

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:30:58 ; Search time 1643.83 seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-23

Perfect score: 29

Sequence: 1 nngaauncuuunnguaagccnangnngn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	62.1	325	4	AA290106	AA290106 vb32e03.r
2	18	62.1	527	226	AQ267968	AQ267968 RPII11-73
3	18	62.1	596	226	AQ267210	AQ267210 RPII11-73
4	17	58.6	681	219	CNS004AD	AL078497 Drosophila
5	17	58.6	943	219	CNS004CW	AL058429 Drosophila
6	17	58.6	985	219	CNS000J5	AL076625 Drosophila
7	17	58.6	988	219	CNS000J7C	AL076711 Drosophila
8	17	58.6	1101	219	CNS000BEM	AL056770 Drosophila
9	17	58.6	1101	219	CNS000CCH	AL058573 Drosophila
10	16	55.2	330	162	BE065429	BE065429 RC3-BT0313
11	16	55.2	494	113	AM240011	AM240011 pTTC.pK0
12	16	55.2	888	169	BF795148	BF795148 602256488
13	15.4	53.1	145	150	BF522395	BF522395 UT-R-C3-L
14	15.4	53.1	272	26	AA0205618	AA0205618 AV205618
15	15.4	53.1	311	224	AA091856	AA091856 HS_2208.A
16	15.4	53.1	316	249	AT262135	AZ7622151 LM0556B19
17	15.4	53.1	397	21	AI535204	AZ7622151 LM0556B19
18	15.4	53.1	408	224	AQ148119	AI535204 UT-R-C3-L
19	15.4	53.1	417	123	AQ065332	AQ148119 HS_3109.A
20	15.4	53.1	533	103	AI914074	AQ065332 HS_2227.B
21	15.4	53.1	554	164	BE173543	AI914074 w474612.x
22	15.4	53.1	591	242	AZ377159	BE173543 RC2-HT0565
23	15.4	53.1	602	236	AQ970351	AZ377159 RPI1-1B22
24	15.4	53.1	604	244	AZ486340	AQ970351 RPI1-23-3
25	15.4	53.1	613	149	BE472495	AZ486340 LM0314J18
26	15.4	53.1	691	164	BE214575	BE472495 UT-M-BH3
27	15.4	53.1	707	241	AZ291529	BE214575 HV_CFB000
28	15.4	53.1	828	220	CNS010MP	AZ291529 RPI1-23-I
29	15.4	53.1	851	221	CNS03KAZ	AL170206 Tetradodon
30	15.4	53.1	1048	221	CNS03KX1	AL247940 Tetradodon
31	15.4	53.1	1086	221	CNS04008	AL247870 Tetradodon
32	15.4	53.1	1665	144	BF120703	AL268649 Tetradodon
33	15	51.7	234	25	AV130117	BF120703 601758286
34	15	51.7	268	25	AV142328	AV130117 AV130117
35	15	51.7	436	189	T997938	AV142328 AV142328
36	15	51.7	501	152	BG370061	T997938 ye67e11.r1
37	15	51.7	584	229	AQ465094	BG370061 LMA061.V8
38	15	51.7	660	230	AQ586492	AQ465094 HS_5084.B
39	15	51.7	915	153	BG390703	AQ586492 RPI1-11-C
40	14.8	51.0	138	144	BF077310	BG390703 602416413
41	14.8	51.0	189	249	AZ767655	BF077310 227157.MA
42	14.8	51.0	248	191	Z28552	AZ767655 LM0567A06
43	14.8	51.0	293	15	AV154534	Z28552 HSB35G061.S
44	14.8	51.0	293	25	AT057753	AV154534 AV154534
45	14.8	51.0	316	147	BF377814	AT057753 TENDU1844
						BF377814 CM2-TN013

ALIGNMENTS

RESULT	1
LOCUS	AA290106
DEFINITION	AA290106 325 bp mRNA EST 14-APR-1997
ACCESSION	U62603.1
VERSION	1
KEYWORDS	IMAGE:750652.5' similar to gb:M16762 Mouse interleukin 2 (MOUSE);, mRNA sequence.
SOURCE	AA290106
ORGANISM	AA290106.1 GI:1936335
REFERENCE	EST.
AUTHORS	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 325) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

rhelsing, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
mml@wustl.edu

FEATURES

Therising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The Mashu-HHMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@waterston.wustl.edu
This clone is available royalty-free through LINL; contact the IMAGC Consortium (info@image.llnl.gov) for further information.
MGI:459636
Putative full length read
vector to vector length is 422
Seq primer:-28m3 rev2 ET from Amersham.

Source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:750652"
 /clone_lib="Soares mouse lymph node NdMLN"
 /sex="male"
 /tissue_type="lymph node"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pF773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer; 15'
 TGTACCAATCTGAAGTGGAGCGCGCGATACCTTTTTTTTTTTTTTTTTTT
 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7n3 vector. RNA provided by Dr. Bertrand Jordan, library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN

Query Match	62.1%	Score 18:	DB 4;	Length 325;
Best Local Similarity	54.2%	Pred. No.	9.5;	
Matches 13; Conservative	5;	Mismatches	6;	Indels 0; Gaps 0;

QY 4

Qy 4 gauncuunnguagcccnang 27
||:|::|:||||| | | |
Db 170 GATTCCTTTTGTAGCCCCAAGG 19

RESULT 2
 A0267968/c
 LOCUS A0267968 527 bp DNA GSS 27-APR-1999
 DEFINITION R0C11-73E11.TU R0C1-11 Homo sapiens genomic clone R0C1-11-73E11,
 DNA sequence.
 ACCESSION A0267968
 VERSION A0267968.1 GI:3795572
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 527)
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C
 Use of human BAC End Sequences for Sequence-Ready Map Building
 Unpublished (1998)
 Other-GSSs: R0C11-73E11.TK
 COMMENT
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA

RESULT 5
CNS00C4W 943 bp DNA GSS 04-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR25E15 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL058429
ACCESSION AL058429.1 GI:4946310
KEYWORDS
SOURCE
ORGANISM
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers

FEATURES
source
1. 943
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR25E15"
/note="end : T7"

BASE COUNT 331 a 166 c 143 g 265 t 38 others
ORIGIN

Query Match 58.6%; Score 17; DB 219; Length 943;
Best Local Similarity 52.2%; Pred. No. 45;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 5 auncuununguaagccnang 27
I: :::: |::|::|::|::|
Db 398 ATCCTTGGGTAGCCGAGTG 420

RESULT 6
CNS00J05 985 bp DNA GSS 03-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR38L15 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL076625
ACCESSION AL076625.1 GI:4956102
KEYWORDS
SOURCE
ORGANISM
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission

JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers

FEATURES
source
1. 985
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR38L15"
/note="end : TET3"

BASE COUNT 323 a 172 c 148 g 303 t 39 others
ORIGIN

Query Match 58.6%; Score 17; DB 219; Length 985;
Best Local Similarity 52.2%; Pred. No. 45;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 5 auncuununguaagccnang 27
I: :::: |::|::|::|::|
Db 462 ATCCTTGGGTAGCCGAGTG 484

RESULT 7
CNS00J0C 988 bp DNA GSS 03-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR39I07 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL076711
ACCESSION AL076711.1 GI:4956289
KEYWORDS
SOURCE
ORGANISM
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

TITLE	'M.J.', Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
JOURNAL	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT	2020263 Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC3-BR0316-071>)
299-011-e01a6t3=1999-12-07&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 330.

```

ATUCBS
source
Location/Qualifiers
1..330
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="pM0316"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 156,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT
100 a 64 c 73 g 93 t

```

Query Match	55.28;	Score 16;	DB 162;	Length 330;
Best Local Similarity	52.48;	Pred. No. 1.3e+02;		
Matches 11; Conservative	5;	Mismatches 5;	Indels 0;	Gaps 0;

```
Qy      5 auncuunnguagcccnang 25
          |::: | | | | | |
Db      184 ATGCTTATGTAGCCCAAG 204
```

RESULT 11
AW240011

Accession	Length	Source	Date
F06789	494 bp	mRNA	27-APR-2000
AW240011			
ptlc;pk001.120	chicken T cell cDNA library	Gallus gallus	cDNA clone
ptlc;pk001.120	5', mRNA sequence.		

ACCESSION	AW240011	GI:6579751
VERSION	AW240011.1	
KEYWORDS	EST.	
SOURCE	chicken.	

ORGANISM
Gallus gallus
Eukaryote; Metazoa; Chordata; Craniata; Euteleostomi
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
I (bases 1 to 494)

AUTHORS	Morgan, R.
TITLE	Chicken T cell ESTs
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robin Morgan

University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341

FEATURES
source
Clones can be ordered online at <http://www.chickest.udel.edu>
Location/Qualifiers
I. 494

```

/db.xref="taxon:9031"
/clone="ptfrc.pk001.120"
/clone.lib="chicken.T.cell.cDNA.library"
/sex="Male and Female T cell"
/cell_type="splenic T cell"
/lab_host="E.coli DH10B"
/note="Vector: pB42AD"

BASE COUNT      151 a       82 c       96 g      143 t       22 others
ORIGIN

Query Match          55.2%   Score 16: DB 113:
Best Local Similarity 60.0%:   Pred No. 1.4e+02:
Matches    12; Conservative     5; Mismatches    3; Indels      0; Gaps      0

OY      4 gauncuuunnguagaaccna 23
         |||: |::: |::| | | | | | |
db      18 gatgcctttrtgtaagcccca 37

```

RESULT	12
BF795148/c	
LOCUS	BF795148.1
DEFINITION	602256488F1 NTH_MGC_85 Homo sapiens cDNA clone IMAGE:4339844 5', mRNA sequence.
ACCESSION	BF795148
VERSION	BF795148.1
KEYWORDS	GI:12100202
SOURCE	EST.
	human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 888)
REFERENCE	NH-MGC http://mgc.nci.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	

Email: cgabps-r@mail.nih.gov
Tissue Procurement: Louis Straud, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLM950 row: p column: 21
High quality sequence stop: 678.

FEATURES	Location/Qualifiers
source	1. .888

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4339844"
/clone_id="NIH_MGC_85"
/tissue_type="lymphoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph, Vector: PCMV-SPORT6, Site_1: NotI;
Site_2: SalI. Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: This is a NIH_MGC Library."

```

BASE COUNT	217 a	217 c	236 g	218 t
ORIGIN				

Query Match	55.2%	Score 16:	DB 169;	Length 888;
Best Local Similarity	50.0%	Pred. No.	1.7e+02;	
Matches 11; Conservative	5;	Mismatches	6;	Indels 0; Gaps 0

```
QY      6 uncuuunnguagcccnang 27
      :|:::|:||||| | | |
Db    817 TCCTTCCGTAAGCCCCAAGTG 796
```

```

RESULT 13
LOCUS BF522395/c 145 bp mRNA EST 11-DEC-2000
DEFINITION UI-R-C3-ta-e-05-0-UI-r2 UI-R-C3 Rattus norvegicus cDNA clone
ACCESSION BF522395
VERSION BF522395.1 GI:11630362
KEYWORDS EST
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 145)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
AUTHORS Contact: Soares, MB
TITLE Program for Rat Gene Discovery and Mapping
JOURNAL University of Iowa
MEDLINE 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
AUTHORS Tel: 319 335 8250
FAX: 319 335 9565
Email: msoares@iuiiue.weeg.uiowa.edu
CDNA library preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LMN (info@image.liml.gov). IMAGE ID-1769254
Seq primer: M13 Forward
FEATURES
Location/Qualifiers
1..145
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C3-ta-e-05-0-UI"
/clone_lib="UI-R-C3"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-C3
library is a subtracted library of a series, ultimately
derived from a mixture of individually tagged normalized
libraries from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day
embryos, after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1
, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-C3) was constructed as follows: PCR amplified
cDNA inserts from UI-R-C2p clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the UI-R-C2p library in the form of single-stranded
circles. The remaining single-stranded circles (subtracted
library) was purified by hydroxyapatite column
chromatography, converted to double-stranded circles and
electroporated into DH10B bacteria (Life Technologies) to
generate the UI-R-C3 library. This procedure has been
previously described (Bonaldo, Lennon and Soares, Genome
Research 6:791-806, 1996)"
BASE COUNT 23 a 39 c 40 g 43 t
ORIGIN
Query Match 53.1%; Score 15.4; DB 150; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

Oy 4 gauncuunungaaaccnang 25
Db 75 GATGCTTTCTAAGCCACAG 54
RESULT 14
LOCUS AV205618/c 272 bp mRNA EST 30-OCT-1999
DEFINITION AV205618 RIKEN full-length enriched, adult male testis Mus musculus
ACCESSION AV205618
VERSION AV205618.1 GI:6146471
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 272)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai,
C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata,
Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H.,
Suzuki, H., Takahashi, F., Tateo, M., Tomimaga, N., Tsunoda, Y.,
Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T.,
Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rict.riken.go.jp,
URL: http://genome.rict.riken.go.jp/
Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki,
Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rict.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..272
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1700081C10"
/clone_lib="RIKEN full-length enriched, adult male testis"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'

```

Tue Oct 2 09:31:45 2001

Search completed: October 2, 2001, 04:38:38
Job time: 4060 sec

BASE COUNT	90 a	52 c	69 g	61 t
ORIGIN				

Best Local Similarity	50.0%;	Pred No. 2.8e+02;							
Matches	11;	Conservative	5;	Mismatches	6;	Indels	0;	Gaps	0;

```

||:|::|:|:|:|:|:|
Db 112 GATGCTTCTGTAAGCTCCAGG 91

```

SULT 15
091856

ACCESSION	
DEFINITION	HS_2208_Al_H1_MK_CIT Approved Human Genomic Sperm Library b Homo sapiens genomic clone Plate=2208 Col=21 Row=0, DNA sequence.
VERSION	W001056

ACCESSION AQ091856

VERSION AQ091856.1 GI:3460767
KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 311)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

EST

JOURNAL
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999).
MEDLINE
99380589

COMMENT Contact: Mahairas GG, Wallace JC,

Tel: (206) 616-3618
Fax: (206) 616-3887

Class: BAC ends
High quality sequence stop: 311.

FEATURES	Location/Qualifiers
source	1. .311

BASE COUNT	134 a	28 c	51 g	95 t	3 others
ORIGIN					

```
Best Local Similarity 50.08; Pred. No. 2.9e+02;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

Db 217 GATGTTTATGTAAGCCCATG 238

Tue Oct 2 09:31:45 2001

us-09-310-844b-23.rst

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:33:13 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-23

Sequence: 1 mngaucununguaagccnangn 29

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 324599 seqs, 94655562 residues

total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	62.1	209	1	US-08-510-032A-8
2	18	62.1	209	3	US-08-688-514-8
3	18	62.1	801	6	5314995-8
4	18	62.1	8491	2	US-08-757-439-1
5	14.4	49.7	1894	4	US-09-004-731-29
6	14.4	49.7	1894	4	US-09-004-731-31
7	14.4	49.7	1894	4	US-09-032-215-3
8	14.4	49.7	1894	4	US-08-749-699-29
9	14.4	49.7	1894	4	US-08-749-699-31
10	14.4	49.7	10409	3	US-08-772-440-33
11	14.4	49.7	21126	1	US-08-008-216-19
12	14.4	49.7	21126	1	US-08-459-569-19
13	14.4	49.7	21126	1	US-08-458-831-19
14	13.8	47.6	2670	1	US-08-121-713D-61
15	13.8	47.6	2670	1	US-08-835-268-61
16	13.8	47.6	2670	2	US-09-060-692-61
17	13.8	47.6	2670	2	US-08-833-391-61
18	13.8	47.6	2670	5	PCT-US94-10151A-61
19	13.4	46.2	2829	3	US-08-851-843A-53
20	13.4	46.2	2829	4	US-08-974-549A-220
21	13.2	45.5	945	4	US-09-434-774-15
22	13.2	45.5	3867	3	US-08-762-428A-5
23	13.2	45.5	4016	3	US-08-762-428A-7
24	13.2	45.5	4884	1	US-07-665-792E-10
25	13	44.8	172	1	US-08-480-910-8
26	13	44.8	300	1	US-08-480-910-2
27	13	44.8	1285	1	US-08-480-910-1

C	28	13	44.8	1285	5	PCT-US95-00052-1	Sequence 1, Appl
	29	12.8	44.1	423	1	US-08-470-179-41	Sequence 41, Appl
	30	12.8	44.1	423	1	US-08-470-179-120	Sequence 120, Appl
	31	12.8	44.1	423	1	US-08-470-179-123	Sequence 123, Appl
	32	12.8	44.1	705	4	US-08-998-416-1136	Sequence 1136, Ap
	33	12.8	44.1	1210	3	US-09-002-298-2	Sequence 2, Appl1
	34	12.8	44.1	1212	3	US-08-188-930-249	Sequence 249, App
	35	12.8	44.1	1803	3	US-08-458-922-2	Sequence 2, Appl1
	36	12.8	44.1	2221	1	US-08-418-782-2	Sequence 2, Appl1
	37	12.8	44.1	2221	1	US-08-228-662-2	Sequence 2, Appl1
	38	12.8	44.1	2221	2	US-08-852-219-2	Sequence 2, Appl1
	39	12.8	44.1	2235	1	US-08-418-782-1	Sequence 1, Appl1
	40	12.8	44.1	2235	1	US-08-228-662-1	Sequence 1, Appl1
	41	12.8	44.1	2235	2	US-08-852-219-1	Sequence 1, Appl1
	42	12.8	44.1	2331	1	US-08-418-782-20	Sequence 20, Appl
	43	12.8	44.1	2331	2	US-08-852-219-20	Sequence 20, Appl
	44	12.8	44.1	2955	2	US-08-867-941-4	Sequence 4, Appl1
	45	12.8	44.1	2955	4	US-09-074-658-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-510-032A-8
; Sequence 8, Application US/08510032A
; Patent No. 5712126

GENERAL INFORMATION:

APPLICANT: Sherman Weissman and Yalindra Prashar

TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSER: Yahwak & Associates
STREET: 25 Skytop Drive

CITY: Trumbull
STATE: Connecticut
COUNTRY: USA

ZIP: 06611

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word 4.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,032A

FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak

REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: Yale

TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951

TELEFAX: (203)268-1951
INFORMATION FOR SEQ. ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 209 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA

US-08-510-032A-8

Query Match 62.1%; Score 18; DB 1; Length 209;
Best local Similarity 54.2%; Pred. No. 0.064;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 gauncuununguaagccnangn 27
||:|||||
DB 36 GATCTTTTGTAAAGCCCTAGGCG 59

RESULT 2
US-08-688-514-8
Sequence 8, Application US/08688514
Patent No. 6010850
GENERAL INFORMATION:
APPLICANT: Shetman Weissman and Yairindra Prashar
TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,514
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: Yale
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-688-514-8

Query Match 62.1%; Score 18; DB 3; Length 209;
Best Local Similarity 54.2%; Pred. NO. 0.064;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

4 gauncuununguaagcccnang 27
||:|::|:||||| | | |
36 GATCTCTTTGTGAAGCCCTAGGCG 59

RESULT 3
5314995-8
Patent No. 5314995
APPLICANT: FELL, HENRY P.; GAYLE, MARGIT A.
TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
BASED FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/468,390
FILING DATE: 22-JAN-1990
SEQ ID NO: 8
LENGTH: 801
5314995-8

Query Match 62.1%; Score 18; DB 6; Length 801;
Best Local Similarity 54.2%; Pred. NO. 0.086;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

4 gauncuununguaagcccnang 27
||:|::|:||||| | | |

Db 653 gatctcttgtgaagccctaggg 676

RESULT 4
US-08-757-439-1/c
Sequence 1, Application US/08757439
Patent No. 5866371
GENERAL INFORMATION:
APPLICANT: BADZIONG, Werner
APPLICANT: HABERMANN, Paul
APPLICANT: MOELLER, Joerg
APPLICANT: ARETZ, Werner
TITLE OF INVENTION: PROCESS FOR USING THE YEAST ADH II
PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
PROTEINS IN HIGH YIELDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,439
FILING DATE: 27-NOV-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19544233.4
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 18748/303/HOCE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8491 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-757-439-1

Query Match 62.1%; Score 18; DB 2; Length 8491;
Best Local Similarity 54.2%; Pred. NO. 0.15;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

4 gauncuununguaagcccnang 27
||:|::|:||||| | | |
Db 6651 GATCTCTTTGTGAAGCCCTAGGCG 6628

RESULT 5
US-09-004-731-29/c
Sequence 29, Application US/09004731
Patent No. 6177258
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:

```

1 ADDRESS: Sheridan Ross P.C.
2 STREET: 1700 Lincoln Street, Suite 3500
3 CITY: Denver
4 STATE: Colorado
5 COUNTRY: USA
6 ZIP: 80203
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12 SOFTWARE: Patentln Release #1.0, Version #1.30
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/09/004,731
16 FILING DATE:
17 CLASSIFICATION:
18
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US/08/749,699
21 FILING DATE:
22
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Connell, Gary J.
25 REGISTRATION NUMBER: 32,020
26 REFERENCE/DOCKET NUMBER: 2618-25-C3
27
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (303) 863-9700
30 TELEFAX: (303) 863-0223
31
32 INFORMATION FOR SEQ ID NO: 29:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 1894 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38
39 MOLECULE TYPE: DNA (genomic)
40
41 FEATURE:
42
43 NAME/KEY: CDS
44 LOCATION: 335..1535
45
46 US-09-004-731-29
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

```

1 CLASSIFICATION:
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US/08/749,699
4 FILING DATE:
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Connell, Gary J.
7 REGISTRATION NUMBER: 32,020
8 REFERENCE/DOCKET NUMBER: 2618-25-C3
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (303) 863-9700
11 TELEFAX: (303) 863-0223
12 INFORMATION FOR SEQ ID NO: 31:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 1894 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: DNA (genomic)
19 US-09-004-731-31
20
21 Query Match 49.7%; Score 14.4; DB 4; Length 1894;
22 Best local similarity 47.6%; Pred. No.15;
23 Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0.
24
25 Oy 5 auncuunnguagcccnang 25
26 1:|::|:|::|:|
27
28 Db 328 ATACTTGGGTAAAGCTCGATG 348
29
30 RESULT 7
31 US-09-032-215-3/c
32 Sequence 3, Application US/09032215
33 Patent No. 6204010
34 GENERAL INFORMATION:
35 APPLICANT: Stiegler, Gary L.
36 APPLICANT: Gaines, Patrick J.
37 TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC
38 TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
39 NUMBER OF SEQUENCES: 50
40 CORRESPONDENCE ADDRESS:
41 ADDRESSEE: Sheridan Ross P.C.
42 STREET: 1700 Lincoln Street, Suite 3500
43 CITY: Denver
44 STATE: Colorado
45 COUNTRY: U.S.A.
46 ZIP: 80203
47
48 COMPUTER READABLE FORM:
49 MEDIUM TYPE: Floppy disk
50 COMPUTER: IBM PC compatible
51 OPERATING SYSTEM: PC-DOS/MS-DOS
52 SOFTWARE: ASCII DOS TEXT
53 CURRENT APPLICATION DATA:
54 APPLICATION NUMBER: US/09/032,215
55 FILING DATE: 27-FEB-1998
56 CLASSIFICATION: 536
57 PRIOR APPLICATION DATA:
58 APPLICATION NUMBER:
59 FILING DATE:
60 ATTORNEY/AGENT INFORMATION:
61 NAME: Connell, Gary J.
62 REGISTRATION NUMBER: 32,020
63 REFERENCE/DOCKET NUMBER: 2618-25-C6
64 TELECOMMUNICATION INFORMATION:
65 TELEPHONE: (303) 863-9700
66 TELEFAX: (303) 863-0223
67 INFORMATION FOR SEQ ID NO: 3:
68 SEQUENCE CHARACTERISTICS:
69 LENGTH: 1894 nucleotides
70 TYPE: nucleic acid
71 STRANDEDNESS: single
72 TOPOLOGY: linear
73 MOLECULE TYPE: cDNA

```



```

1 PatentNO.: 5366887
2 GENERAL INFORMATION:
3 APPLICANT: Slightom, Jerry L.
4 APPLICANT: Tepfer, David A.
5 TITLE OF INVENTION: R1 T-DNA Promoters
6 NUMBER OF SEQUENCES: 19
7 CORRESPONDENCE ADDRESS:
8 ADDRESS: GRAY, CARL, AMES & FRYE
9 STREET: 401 B Street, Suite 1700
10 CITY: San Diego
11 STATE: California
12 COUNTRY: USA
13 ZIP: 92101-4297
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: PatentIn Release #1.0, Version #
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/008,216
21 FILING DATE: 25-JAN-1993
22 CLASSIFICATION: 800
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 06/725,368
25 FILING DATE: 22-APR-1985
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Barnhorst, Marnie W.
28 REGISTRATION NUMBER: 36,740
29 REFERENCE/DOCKET NUMBER: PI1020U1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (619) 699-2700
32 TELEFAX: (619) 236-1048
33 INFORMATION FOR SEQ ID NO.: 19:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 21126 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 MOLECULE TYPE: DNA (genomic)
40 HYPOTHEetical: NO
41 ORIGINAL SOURCE:
42 ORGANISM: Agrobacterium rhizogenes
43 STRAIN: STRAIN A4
44 IMMEDIATE SOURCE:
45 LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
46 CLONE: CLONE 7
47 FEATURE:
48 NAME/KEY: misc_feature
49 LOCATION: complement (937..2262)
50 OTHER INFORMATION: /label= ORF1SUBSEQUENCE
51 FEATURE:
52 NAME/KEY: misc_feature
53 LOCATION: complement (2649..3458)
54 OTHER INFORMATION: /label= ORF2SUBSEQUENCE
55 FEATURE:
56 NAME/KEY: misc_feature
57 LOCATION: 3726..4799
58 OTHER INFORMATION: /label= ORF3SUBSEQUENCE
59 FEATURE:
60 NAME/KEY: misc_feature
61 LOCATION: complement (4041..4400)
62 OTHER INFORMATION: /label= ORF4SUBSEQUENCE
63 FEATURE:
64 NAME/KEY: misc_feature
65 LOCATION: complement (4607..4918)
66 OTHER INFORMATION: /label= ORF5SUBSEQUENCE
67 FEATURE:
68 NAME/KEY: misc_feature
69 LOCATION: 5143..6216
70 OTHER INFORMATION: /label= ORF6SUBSEQUENCE
71 FEATURE:
72 NAME/KEY: misc_feature
73 LOCATION: complement (5071..5643)

```

OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:--
NAME/KEY: misc_feature
LOCATION: 6609..888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12466..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (16517..17545)
OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF16SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (18177..18743)
OTHER INFORMATION: /label= ORF17SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (19031..19390)
OTHER INFORMATION: /label= ORF18SUBSEQUENCE
US-08-008-216-19
Query Match 49.7%; Score 14.4; DB 1; Length 21126;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 5 auncuunungaaagcccnang 25
Db 20133 ATTCTTGGGTAAGACGCAAG 20153
RESULT 12
US-08-459-569-19
Sequence 19, Application US/08459569
Patent No. 5543501
GENERAL INFORMATION:
APPLICANT: Slightom, Jerry L.
TITLE OF INVENTION: R1 T-DNA Promoters
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: GRAY, CARY, AMES & FRYE
STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-4297

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459, 569
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008, 216
FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 06/725, 368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barnhorst, Marjorie W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: STRAIN A4
IMMEDIATE SOURCE:
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
CLONE: CLONE 7
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (937..2262)
OTHER INFORMATION: /label= ORF1SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (2649..3458)
OTHER INFORMATION: /label= ORF2SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3726..4799
OTHER INFORMATION: /label= ORF3SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (4041..4400)
OTHER INFORMATION: /label= ORF4SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (4607..4918)
OTHER INFORMATION: /label= ORF5SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6609..8888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature

LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12466..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (16517..17545)
OTHER INFORMATION: /label= ORF15SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF16SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (18177..18743)
OTHER INFORMATION: /label= ORF17SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (19031..19390)
OTHER INFORMATION: /label= ORF18SUBSEQUENC
US-08-459-569-19

Query Match 49.7%; Score 14.4; DB 1; Length 21126;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 5 auncununguaagcccnang 25
Db 20133 ATTCTTCGCTAGACCGAAG 20153

RESULT 13
US-08-458-831-19
Sequence 19, Application US/08458831
Patent No. 5824866
GENERAL INFORMATION:
APPLICANT: Slightom, Jerry L.
TITLE OF INVENTION: Tefter, David A.
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: GRAY, CARY, AMES & FRYE
STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-4297
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,831
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,216

FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 06/725,368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barnhorst, Marilee W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: STRAIN A4
IMMEDIATE SOURCE:
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
CLONE: CLONE 7
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (937..2262)
OTHER INFORMATION: /label= ORF1SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (2649..3458)
OTHER INFORMATION: /label= ORF2SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3726..4799
OTHER INFORMATION: /label= ORF3SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (4041..4400)
OTHER INFORMATION: /label= ORF4SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6609..8888
OTHER INFORMATION: /label= ORF8SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12466..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENC
FEATURE:

```
NAME/KEY: misc_feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16517..17545
OTHER INFORMATION: /label= ORF15SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 17737..18189
OTHER INFORMATION: /label= ORF16SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 18177..18743
OTHER INFORMATION: /label= ORF17SUBSEQUENC
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19031..19390
OTHER INFORMATION: /label= ORF18SUBSEQUENC
US-08-458-831-19
```

```
Query Match 49.7%; Score 14.4; DB 1; Length 21126;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

OY 5 auncuununguaagcccnang 25

Db 20133 ATCTTTCGTAAGACCGAAG 20153

```
RESULT 14
US-08-121-713D-61
Sequence 61, Application US/08121713D
Patent No. 5639856
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Mathes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713D
FILING DATE: 13-SEP-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
```

```
LENGTH: 2670 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 268..2439
US-08-121-713D-61
```

```
Query Match 47.6%; Score 13.8; DB 1; Length 2670;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

OY 4 gauncuununguaagcccnang 25

Db 2416 GAGATTTCGTAAGCCCATG 2437

```
RESULT 15
US-08-835-268-61
Sequence 61, Application US/08835268
Patent No. 5807826
GENERAL INFORMATION:
```

```
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Mathes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 2670 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 268..2439
US-08-835-268-61
Query Match 47.6%; Score 13.8; DB 1; Length 2670;
```


Best Local Similarity 50.0%; Pred. No. 37;
Matches 11; Conservative 4; Mismatches

7; Indels 0; Gaps 0;

QY 4 gauncuununguaagccnang 25
|| ::::|:|||||
Db 2416 GAGATTTCGTAAGCCCAATG 2437

Search completed: October 2, 2001, 05:01:46
Job time: 5313 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:31:53 ; Search time 1315.38 Seconds

(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-23

Sequence: 1 mnngauncuunnguaagcccnangngn 29

Scoring table: IDENTITY_NUC

Searched: 1344157 seqs, 7733874588 residues

2688314

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl : *

```

1:  gb_emb1.*
2:  gb_ba2.*
3:  gb_ba3.*
4:  gb_in1.*
5:  gb_in2.*
6:  gb_in3.*
7:  gb_ov.*
8:  gb_ov.*
9:  gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_p12.*
13: gb_p13.*
14: gb_p13.*
15: gb_p14.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_hgo_hum.*
20: em_hgo_in1.*
21: em_hgo_in2.*
22: em_hgt_hum1.*
23: em_hgt_hum2.*
24: em_hgt_hum3.*
25: em_hgt_hum4.*
26: em_hgt_hum5.*
27: em_hgt_hum6.*
28: em_hgt_hum7.*
29: em_hgt_hum8.*
30: em_hgt_in1.*
31: em_hgt_in2.*
32: em_hgt_other.*
33: em_hgt_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in1.*
42: em_om.*
43: em_or.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	62.1	209	10	182323	Sequence 8
2	18	62.1	292	54	G06364	human SRS W
3	18	62.1	664	7	MAU79187	U79187 Mitochondria an
4	18	62.1	689	9	A17717	Plasmid pTG
5	18	62.1	698	9	A02159	Synthetic g
6	18	62.1	698	9	A04681	Synthetic g
7	18	62.1	698	9	A14089	Synthetic D
8	18	62.1	722	10	I03416	Sequence 2

```

9      18      62.1      730      91      GIBL12      K03174 Ape (gibbon
10     18      62.1      740      95      RATL12      M22899 Rat interle
11     18      62.1      750      10     I01391      I01391 Sequence 3
12     18      62.1      764      7      BOVIL2      M12791 Bovine inte
13     18      62.1      769      10     E01460      E01460 CDNA encodi
14     18      62.1      769      10     E02201      E02201 DNA encodin
15     18      62.1      773      7      CFU28141      U28141 Canis famli
16     18      62.1      779      7      CATIL2A      L19402 Felis catu
17     18      62.1      782      10     I01101      I01101 Sequence 2
18     18      62.1      784      9      A06879      A06879 Artificial
19     18      62.1      784      9      A14095      A14095 Synthetic D
20     18      62.1      784      9      A17732      A17732 Plasmid pTG
21     18      62.1      784      10     E00250      E00250 DNA coding
22     18      62.1      784      10     I08366      I08366 Sequence 2
23     18      62.1      788      10     I07944      I07944 Sequence 1
24     18      62.1      790      10     I04498      I04498 Sequence 1
25     18      62.1      790      10     I08377      I08377 Sequence 1
26     18      62.1      791      8      MGA7463      A1007463 Meleagris
27     18      62.1      794      10     E00210      E00210 CDNA encodi
28     18      62.1      794      10     E00211      E00211 CDNA encodi
29     18      62.1      794      10     E00215      E00215 CDNA encodi
30     18      62.1      794      10     E02011      E02011 CDNA encodi
31     18      62.1      794      10     E02018      E02018 CDNA encodi
32     18      62.1      795      10     E00214      E00214 CDNA encodi
33     18      62.1      795      10     E00267      E00267 DNA sequenc
34     18      62.1      798      8      AF033563      AF033563 Gallus ga
35     18      62.1      798      8      AF017645      AF017645 Gallus ga
36     18      62.1      800      10     E59473      E59473 Novel cyto
37     18      62.1      801      9      A04961      A04961 Artificial
38     18      62.1      801      9      A06759      A06759 Artificial
39     18      62.1      801      10     E00336      E00336 Human Inter
40     18      62.1      801      10     I01197      I01197 Sequence 1
41     18      62.1      801      10     I04318      I04318 Sequence 3
42     18      62.1      801      93      HSIL02      V00564 Human mRNA
43     18      62.1      810      10     E00978      E00978 CDNA sequen
44     18      62.1      810      10     I01387      I01387 Sequence 1
45     18      62.1      810      10     I01387      I01387 Sequence 1

```

ALIGNMENTS

```

RESULT 1
LOCUS 182323 209 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 8 from patent US 5712126.
ACCESSION 182323
VERSION 182323.1 GI:3210620
KEYWORDS
SOURCE
ORIGIN
ORGANISM
Unknown.
Unclassified.
1 (bases 1 to 209)
REFERENCE
Weissman,S.M. and Prashar,Y.
Analysis of gene expression by display of 3-end restriction
fragments of cDNA
Patent: US 5712126-A 8 27-JAN-1998;
FEATURES
source
1..209
location/Qualifiers
BASE COUNT 56 a 30 c 33 g 90 t
ORIGIN

```

```

Query Match 62.1%; Score 18; DB 10; Length 209;
Best Local Similarity 54.2%; Pred. No. 1.5;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 4 gauncuununguaagcccnangng 27
||:|||||:|||||:|
Db 36 GATTCCTTTTGTAAAGCCCTAGGG 59

```

```

RESULT 2
LOCUS G06364 292 bp DNA STS 19-OCT-1995
DEFINITION human STS WI-7035.
ACCESSION G06364
VERSION G06364.1 GI:859609
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene
collection.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivora; Hominoidea; Hominidae; Homo.
REFERENCE
1 (bases 1 to 292)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically
Mapped ESTs
JOURNAL Unpublished (1995)
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: TAATTAAGTCCTCCACCTTAAC
Primer B: ATTTCGATTAATTAAGTGAACCA
STS size: 200
PCR Profile:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pm
dNTPs: each 4 mM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 9.3

```

```

FEATURES
source
1..292
location/Qualifiers
STS
primer_bind 1..25
complement(176..200)
BASE COUNT 92 a 32 c 32 g 124 t 12 others
ORIGIN

```

```

Query Match 62.1%; Score 18; DB 54; Length 292;
Best Local Similarity 54.2%; Pred. No. 1.6;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 4 gauncuununguaagcccnangng 27
||:|||||:|||||:|
Db 144 GATTCCTTTTGTAAAGCCCTAGGG 167

```

```

RESULT 3
MAU79187

```

LOCUS MAU79187 664 bp mRNA MAM 03-APR-1998
 DEFINITION Mirounga angustirostris interleukin 2 precursor mRNA, complete cds.
 ACCESSION U79187
 VERSION U79187.1 GI:3015531
 KEYWORDS
 SOURCE northern elephant seal.
 ORGANISM Mirounga angustirostris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.

REFERENCE 1 (bases 1 to 664)
 AUTHORS Shoda, L.K.M., Brown, W.C. and Rice-Ficht, A.C.
 TITLE Sequence and characterization of phocine interleukin 2
 JOURNAL J. Wildl. Dis. 34 (1), 81-90 (1998)
 MEDLINE 98136706
 REFERENCE 2 (bases 1 to 664)
 AUTHORS Shoda, L.K.M., Brown, W.C. and Rice-Ficht, A.C.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-1996) Veterinary Microbiology & Pathology,
 Washington State University, P.O. Box 647040, Pullman, WA
 99164-7040, USA

FEATURES
 source location/Qualifiers
 1..664
 /organism="Mirounga angustirostris"
 /db_xref="taxon:9716"
 sig_peptide 24..83
 CDS 24..488
 /codon_start=1
 /product="interleukin 2 precursor"
 /protein_id="AAC12258.1"
 /db_xref="GI:3015532"
 /translation="MCKKQLLSLCSIALSLVIVANSAPPTSSSTKEQOOLLEQLLDRLRL
 LNVNNEDEPKLSRLTFREYVPRKATETLHQAEELKPLEEVLVLAOSKPFHLTD
 IKELMSNINVTILKLGSETRKCEYDEDTATITPELNMKMTFCOSIIFSTL"

mat_peptide 84..485
 /product="interleukin 2"
 BASE COUNT 219 a 131 c 106 g 208 t
 ORIGIN

Query Match 62.1%; Score 18; DB 7; Length 664;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagcccnang 27
 ||::: ||||| |||
 Db 635 GATTCTTTTGTAAAGCCCTAGGCG 658

RESULT 4
 LOCUS A17717 689 bp mRNA PAT 17-MAY-1994
 DEFINITION Plasmid pTG 26 mRNA for human IL-2.
 ACCESSION A17717
 VERSION A17717.1 GI:513949
 KEYWORDS
 SOURCE unidentified.
 ORGANISM Plasmid unidentified
 unclassified.
 1 (bases 1 to 689)
 AUTHORS Kieny, M.P., Sondermeyer, P. and Lecocq, J.P.
 TITLE Expression of human IL-2 in mammalian cells by a recombinant pox
 virus
 JOURNAL Patent: EP 0206939-A 1 30-DEC-1986;
 TRANSGENE S.A
 FEATURES
 source location/Qualifiers
 1..689
 /organism="unidentified"
 /plasmid="pTG26"
 /db_xref="taxon:32644"
 1..384
 /partial
 /codon_start=1
 /product="human interleukin 2"

/protein_id="CAA01347.1"
 /db_xref="GI:513950"
 /translation="TKKQLOLEPHLLDQMLINGINNKPKLRMLTFKRYMKKA
 TELKHQCLEBEELKPLEEVLVLAOSKPFHLRPRDLISNINVTIVLELKGSETTFMCEYA
 DETATITPELNMKMTFCOSIISTL"

BASE COUNT 251 a 117 c 99 g 222 t
 ORIGIN

Query Match 62.1%; Score 18; DB 9; Length 689;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagcccnang 27
 ||: ||:: ||||| |||
 Db 537 GATTCTTTTGTAAAGCCCTAGGCG 551

RESULT 5
 LOCUS A02159 698 bp DNA PAT 21-MAY-1993
 DEFINITION Synthetic gene for interleukin 2 (IL-2) (partial).
 ACCESSION A02159
 VERSION A02159.1 GI:412310
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 698)
 AUTHORS
 TITLE VECTOR FOR THE EXPRESSION IN YEASTS OF INTERLEUKINE-2,
 YEASTS AND METHOD FOR PREPARING INTERLEUKINE-2
 JOURNAL Patent: WO 8503723-A 6 29-AUG-1985;
 FEATURES
 source location/Qualifiers
 1..698
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 10..393
 /partial
 /codon_start=1
 /transl_table=11
 /product="interleukin 2 (IL-2)"
 /protein_id="CAA00227.1"
 /db_xref="GI:412311"
 /translation="TKKQLOLEPHLLDQMLINGINNKPKLRMLTFKRYMKKA
 TELKHQCLEBEELKPLEEVLVLAOSKPFHLRPRDLISNINVTIVLELKGSETTFMCEYA
 DETATITPELNMKMTFCOSIISTL"

BASE COUNT 251 a 117 c 107 g 222 t
 ORIGIN

Query Match 62.1%; Score 18; DB 9; Length 698;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagcccnang 27
 ||: ||:: ||||| |||
 Db 537 GATTCTTTTGTAAAGCCCTAGGCG 550

RESULT 6
 LOCUS A04681 698 bp DNA PAT 24-MAY-1993
 DEFINITION Synthetic gene for interleukin 2 (IL-2) (partial).
 ACCESSION A04681
 VERSION A04681.1 GI:412434
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 698)
 AUTHORS Lemoine, Y., Sondermeyer, P., Loison, G., Aigle, M. and Lecocq, J.P.
 TITLE Yeast-expression vectors for interleukin-2, transformed yeasts and

Query Match 62.1%; Score 18; DB 91; Length 730;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguagcccnangng 27
 ||:|::|:||||| | | |
 Db 630 GATTCCTTTTGTAAAGCCCTAGGG 653

RESULT 10

LOCUS RATTIL2 740 bp mRNA ROD 27-APR-1993
 DEFINITION Rat interleukin 2 mRNA, complete cds.
 ACCESSION M22899
 VERSION M22899.1 GI:204909
 KEYWORDS Interleukin 2
 SOURCE Rat activated T-lymphocytes, cDNA to mRNA, clone pIL-2.8.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 740)
 McKnight,A.J., Mason,D.W. and Barclay,A.N.
 Sequence of rat interleukin 2 and anomalous binding of a mouse
 interleukin 2 cDNA probe to rat MHC class II-associated invariant
 chain mRNA
 JOURNAL Immunogenetics 30, 145-147 (1989)
 MEDLINE 89339608
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
 by A.J. McKnight, 08-MAR-1989.
 FEATURES Location/Qualifiers
 source 1..740
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 15..482
 /note="interleukin 2 precursor"
 /codon_start=1
 /protein_id="AAA41427.1"
 /db_xref="GI:204910"
 /translation="MYSMLACVALTLVLYNSAPTSSPAKTOOHLLEQLLDLOVL
 LKGINVKNLKIPLMLTFEYLPKQATELKHLOCLENEGALQRLVDLTQSSPFLHD
 AGNFTSNIRVYVTLKGSSEKFECDPEPATVVEFLRMVAICOSIISTMTQ"

sig_peptide 15..74
 /note="interleukin 2 signal peptide"
 mat_peptide 75..479
 /note="interleukin 2"
 polyA_signal 713..718
 /note="putative"
 BASE COUNT 228 a 141 c 143 g 228 t
 ORIGIN

Query Match 62.1%; Score 18; DB 95; Length 740;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguagcccnangng 27
 ||:|::|:||||| | | |
 Db 615 GATTCCTTTTGTAAAGCCCAAGG 638

RESULT 11

LOCUS I01391 750 bp ss-DNA PAT 21-MAY-1993
 DEFINITION Sequence 3 from Patent US 4851512.
 ACCESSION I01391
 VERSION I01391.1 GI:270228
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 750)
 AUTHORS Miyaji,H. and Itoh,S.

TITLE Novel human interleukin-2 polypeptide derivative
 JOURNAL Patent: US 4851512-A 3 25-JUL-1989;
 KYOWA HAKKO Kogyo Co., Ltd.;
 Tokyo, Jp;

FEATURES Location/Qualifiers
 source 1..750
 /organism="unknown"

BASE COUNT 276 a 131 c 104 g 239 t
 ORIGIN

Query Match 62.1%; Score 18; DB 10; Length 750;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguagcccnangng 27
 ||:|::|:||||| | | |
 Db 591 GATTCCTTTTGTAAAGCCCTAGGG 614

RESULT 12

LOCUS BOVIL2 764 bp mRNA MAM 27-APR-1993
 DEFINITION Bovine interleukin 2 (IL-2) mRNA, complete cds.
 ACCESSION M12791
 VERSION M12791.1 GI:163204
 KEYWORDS Interleukin 2.
 SOURCE Bovine lymph node, cDNA to mRNA, clone pIL-2-4.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 764)
 Cerretti,D.P., McKereghan,K., Larsen,A., Cantrell,M.A.,
 Anderson,D., Gillis,S., Cosman,D. and Baker,P.E.
 Cloning, sequence, and expression of bovine interleukin 2
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83, 3223-3227 (1986)
 MEDLINE 86205869

COMMENT Draft entry and clean copy sequence for [1] kindly provided by
 D.Cerretti, 12-AUG-1986.
 There is probably only one copy of the interleukin 2 gene in the
 bovine genome.

FEATURES Location/Qualifiers
 source 1..764
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /tissue_type="lymph"

mrna <1..764
 /gene="IL-2"
 1..764
 /gene="IL-2"
 18..77
 /gene="IL-2"
 18..485
 /gene="IL-2"

sig_peptide 18..485
 CDS /note="prepeptide"
 /codon_start=1
 /product="interleukin 2"
 /protein_id="AAA30586.1"
 /db_xref="GI:163205"

/translation="MYKIQLSCIALTALVANGAPTSSSGTNGTKEVKSLLDLOVL
 LEKVNPKNTKISRMHTDPFYKVNATETKIKLLELKITLFEVLNAPSKNLP
 EIKDSMDNIRKIVLELQSGSEFTCEYDQATVNAVDFLNKMTTFOSISTMT"
 78..482
 /gene="IL-2"
 /product="interleukin 2"
 257 a 133 c 123 g 251 t

BASE COUNT 257 a 133 c 123 g 251 t
 ORIGIN 80 bp upstream of HglaI site.

Query Match 62.1%; Score 18; DB 7; Length 764;
 Best Local Similarity 54.2%; Pred. No. 1.9;

Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncununguaagcccnangng 27
||:|::|:||||| | | |
Db 622 GATCTTTTGTGAAGCCTACGGG 645

RESULT 13
E01460 769 bp RNA PAT 29-SEP-1997
LOCUS E01460
DEFINITION cDNA encoding human interleukin 2.
ACCESSION E01460
VERSION E01460.1 GI:2169716
KEYWORDS JP 1987265992-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Onomichi, K., Eto, Y. and Shibai, H.
TITLE PRODUCTION OF VALUABLE SUBSTANCE USING HUMAN CELL
JOURNAL Patent: JP 1987265992-A 1 18-NOV-1987;
AJINOMOTO CO INC
COMMENT OS Human
PN JP 1987265992-A/1
PD 18-NOV-1987
PF 12-MAY-1986 JP 1986107981
PI ONOMICHI KAZUYA, ETO YUZURU, SHIBAI HIROSHIRO PC
C12P21/00,C12N15/00,C12P21/02,(C12P21/02,C12R1:91); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone-PSDI;
FH Key Location/Qualifiers
FT 5'UTR 1..2
FT CDS 3..507
FT 3'UTR
FT 508..814
FT mat_peptide 3..504
FT /product='human interleukin 2'
FT /location/Qualifiers
FT 1..769
FT /organism='Homo sapiens'
FT /db_xref='taxon:9606'
BASE COUNT 280 a 132 c 112 g 245 t
RIGIN

Query Match 62.1%; Score 18; DB 10; Length 769;
Best Local Similarity 54.2%; Pred. No. 1.9;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncununguaagcccnangng 27
||:|::|:||||| | | |
Db 608 GATCTTTTGTGAAGCCTACGGG 631

RESULT 14
E02201 769 bp RNA PAT 29-SEP-1997
LOCUS E02201
DEFINITION DNA encoding interleukin2 (IL-2).
ACCESSION E02201
VERSION E02201.1 GI:2170439
KEYWORDS JP 1990009388-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Murata, M., Eto, Y. and Shibai, H.
TITLE PRODUCTION OF PHYSIOLOGICALLY ACTIVE PROTEIN

JOURNAL Patent: JP 1990009388-A 2 12-JAN-1990;
AJINOMOTO CO INC
COMMENT OS Homo sapiens
PN JP 1990009388-A/2
PD 12-JAN-1990
PF 08-JUL-1988 JP 1988170142
PR 09-MAR-1988 JP 88P 55270
PI MURATA MASAHIRO, ETO YUZURU, SHIBAI HIROSHIRO PC
C12P21/00,C12N15/12,C12N15/26,C12P21/02,(C12P21/02,C12R1:91); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: library-PSD(I)/IL-2;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FT 5'UTR 1..2
FT CDS 3..464
FT /product='interleukin2'
FT /note='IL-2'
FT mat_peptide 3..461
FT /product='interleukin2'
FT /note='IL-2'
FT 465..769
FT 3'UTR
FT /location/Qualifiers
FT 1..769
FT /organism='Homo sapiens'
FT /db_xref='taxon:9606'
BASE COUNT 280 a 132 c 112 g 245 t
ORIGIN

Query Match 62.1%; Score 18; DB 10; Length 769;
Best Local Similarity 54.2%; Pred. No. 1.9;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncununguaagcccnangng 27
||:|::|:||||| | | |
Db 608 GATCTTTTGTGAAGCCTACGGG 631

RESULT 15
CFU28141 773 bp mRNA MAM 29-JUN-1995
LOCUS CFU28141
DEFINITION Canis familiaris interleukin-2 mRNA, complete cds.
ACCESSION U28141
VERSION U28141.1 GI:881935
KEYWORDS dog.
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 773)
AUTHORS Somberg, R.L., Tipold, A., Henthorn, P.S. and Felsburg, P.J.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 773)
AUTHORS Somberg, R.L.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-1995) Richard L. Somberg, Clinical Studies, PA
University of Pennsylvania, 3850 Spruce Street, Philadelphia, PA
19104, USA
FEATURES
source 1..773
/organism='Canis familiaris'
/db_xref='taxon:9615'
39..506
/codon_start=1
/product='interleukin-2'
/protein_id='AA68969.1'
/db_xref='GI:881936'
/translation='MWMQLSCIALTLVANSAPITSSSTETEDQOMQLDLQQL
LNGVNVENPQLSRMLTFKRYTPKATFTTHLQCLAEILKNLEEVGLPQSKNVHLT

polyA_site 773
 BASE COUNT 265 a 145 c 118 g 245 t
 ORIGIN

Query Match 62.1%; Score 18; DB 7; Length 773;
 Best Local Similarity 54.2%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 gauncuuuunguaagccnangng 27
 ||:|::|:||||| | | |
 Db 643 GATTCTTTTGTAAAGCCCTAGGGG 666

Search completed: October 2, 2001, 05:00:40
 Job time: 5327 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:00:28 ; Search time 122.48 Seconds

(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844b-23

Perfect score: 29

Sequence: 1 nmgauuncuuununguaagcccnangnn 29

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	62.1	29	21	AAA70827
2	18	62.1	29	21	AAA70828
3	18	62.1	29	21	AAA70829
4	18	62.1	29	21	AAA70830
5	18	62.1	42	21	AAA71113
6	18	62.1	42	21	AAA71114
7	18	62.1	42	21	AAA71115
8	18	62.1	42	21	AAA71116
9	18	62.1	42	21	AAA71118
10	18	62.1	42	21	AAA71119
11	18	62.1	42	21	AAA71120

12	18	62.1	42	21	AAA71121	Molecular interact
13	18	62.1	42	21	AAA71123	Molecular interact
14	18	62.1	42	21	AAA71124	Molecular interact
15	18	62.1	42	21	AAA71126	Molecular interact
16	18	62.1	42	21	AAA71127	Molecular interact
17	18	62.1	42	21	AAA71128	Molecular interact
18	18	62.1	42	21	AAA71129	Molecular interact
19	18	62.1	42	21	AAA71131	Molecular interact
20	18	62.1	42	21	AAA71132	Molecular interact
21	18	62.1	44	21	AAA71112	Molecular interact
22	18	62.1	44	21	AAA71125	Molecular interact
23	18	62.1	44	21	AAA71133	Molecular interact
24	18	62.1	45	21	AAA70824	Molecular interact
25	18	62.1	45	21	AAA70825	Molecular interact
26	18	62.1	45	21	AAA70826	Molecular interact
27	18	62.1	46	21	AAA71085	Molecular interact
28	18	62.1	46	21	AAA71087	Molecular interact
29	18	62.1	46	21	AAA71088	Molecular interact
30	18	62.1	46	21	AAA71089	Molecular interact
31	18	62.1	46	21	AAA71090	Molecular interact
32	18	62.1	46	21	AAA71093	Molecular interact
33	18	62.1	46	21	AAA71094	Molecular interact
34	18	62.1	46	21	AAA71095	Molecular interact
35	18	62.1	46	21	AAA71096	Molecular interact
36	18	62.1	46	21	AAA71099	Molecular interact
37	18	62.1	46	21	AAA71100	Molecular interact
38	18	62.1	46	21	AAA71103	Molecular interact
39	18	62.1	46	21	AAA71104	Molecular interact
40	18	62.1	46	21	AAA71105	Molecular interact
41	18	62.1	46	21	AAA71106	Molecular interact
42	18	62.1	46	21	AAA71107	Molecular interact
43	18	62.1	46	21	AAA71109	Molecular interact
44	18	62.1	46	21	AAA71110	Molecular interact
45	18	62.1	46	21	AAA71111	Molecular interact

ALIGNMENTS

RESULT 1
ID AAA70827 standard: RNA: 29 BP.
AC AAA70827;
XX
XX 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #27.
XX
XX
XX Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Synthetic.
XX
XX WO9958947-A2.
XX
XX 18-NOV-1999.
XX
XX
XX 12-MAY-1999; 99WO-US10361.
XX
XX 12-MAY-1998; 98US-0076404.
XX
XX 12-MAY-1998; 98US-0085092.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Eckert DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX Hofstadler S, McNeil J;
XX WPI: 2000-086439/07.
XX
XX Identifying compounds which modulate activity of target biomolecules,
XX used to provide compounds which can be used as pharmacological,
XX agricultural and industrial compounds -

PS Claim 235; Page 235; 405pp; English.

XX This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAUCUGUUUACGAAAUUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

CC Sequence 29 BP; 4 A; 4 C; 5 G; 5 U; 11 other;

Query Match 62.1%; Score 18; DB 21; Length 29;

Best Local Similarity 100.0%; Pred. No. 0.35;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 gauncuunungaaagcccnangng 27

||||| ||||||| |||||||

Db 4 gauncuunungaaagcccnangng 27

RESULT 2

AAA70828 ID AAA70828 standard; RNA; 29 BP.

XX AAA70828;

XX 27-APR-2001 (first entry)

XX Molecular interaction site RNA #28.

XX Modulator: identification; molecular interaction; virtual library; ss.

XX Homo sapiens.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX Hofstadler S, McNeil J;

XX WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.

XX This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAUCUGUUUACGAAAUUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

CC Sequence 29 BP; 5 A; 5 C; 7 G; 12 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;

Best Local Similarity 75.0%; Pred. No. 0.35;

Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuunungaaagcccnangng 27

||||| ||||||| |||

Db 4 gauncuunungaaagcccnaggg 27

RESULT 3

AAA70829 ID AAA70829 standard; RNA; 29 BP.

XX AAA70829;

XX 27-APR-2001 (first entry)

XX Molecular interaction site RNA #29.

XX Modulator: identification; molecular interaction; virtual library; ss.

XX Mus sp.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX Hofstadler S, McNeil J;

XX WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405bp; English.

XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCUUGACACAAAUAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;

Best Local Similarity 75.0%; Pred. No. 0.35;

Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguagcccaagng 27

DB 4 gauncuununguagcccaagng 27

RESULT 4

AAA70830

ID AAA70830 standard; RNA: 29 BP.

XX AC AAA70830;

XX DT 27-APR-2001 (first entry)

XX Molecular interaction site RNA #30.

XX Modulator; Identification; molecular interaction; virtual library; ss.

XX OS Rattus sp.

XX PN WO9958947-A2.

XX PD 18-NOV-1999.

XX PF 12-MAY-1999; 99WO-US10361.

XX PR 12-MAY-1998; 98US-0076404.

XX PR 12-MAY-1998; 98US-0085092.

XX PA (ISIS-) ISIS PHARM INC.

XX PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX PI Hostadler S, McNeil J;

XX DR WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,

XX used to provide compounds which can be used as pharmacological,

XX agricultural and industrial compounds -

PS Claim 235; Page 235; 405bp; English.

XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCUUGACACAAAUAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;

Best Local Similarity 75.0%; Pred. No. 0.35;

Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguagcccaagng 27

DB 4 gauncuununguagcccaagng 27

RESULT 5

AAA71113

ID AAA71113 standard; RNA: 42 BP.

XX AC AAA71113;

XX DT 27-APR-2001 (first entry)

XX Molecular interaction site RNA #189.

XX Modulator; Identification; molecular interaction; virtual library; ss.

XX OS Unidentified.

XX PN WO9958947-A2.

XX PD 18-NOV-1999.

XX PF 12-MAY-1999; 99WO-US10361.

XX PR 12-MAY-1998; 98US-0076404.

XX PR 12-MAY-1998; 98US-0085092.

XX PA (ISIS-) ISIS PHARM INC.

XX PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX PI Hostadler S, McNeil J;

XX DR WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,

XX used to provide compounds which can be used as pharmacological,

XX agricultural and industrial compounds -

PS Example 7; Figure 122; 405bp; English.

XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUCUAGUUACACAAAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;

Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagcccaangng 27
DB 7 gauncuununguaagcccaaggg 30
||||| ||||| |||

RESULT 8

AAA71116
ID AAA71116 standard; RNA; 42 BP.

XX
AC AAA71116;

XX
DT 27-APR-2001 (first entry)

XX
KW Molecular interaction site RNA #192.

XX
DE Modulator; Identification; molecular interaction; virtual library; ss.

XX
OS Unidentified.

XX
PN WO9558947-A2.

XX
PD 18-NOV-1999.

XX
PF 12-MAY-1999; 99WO-US10361.

XX
PR 12-MAY-1998; 98US-0076404.

XX
PR 12-MAY-1998; 98US-0085092.

XX
PA (ISIS-) ISIS PHARM INC.

XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hostadler S, McNeil J;

XX
DR WPI: 2000-086439/07.

XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405bp; English.

XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUCUAGUUACACAAAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;

Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagcccaangng 27
DB 7 gauncuununguaagcccaaggg 30
||||| ||||| |||

RESULT 9

AAA71118
ID AAA71118 standard; DNA; 42 BP.

XX
AC AAA71118;

XX
DT 27-APR-2001 (first entry)

XX
KW Molecular interaction site DNA #124.

XX
DE Modulator; Identification; molecular interaction; virtual library; ss.

XX
OS Unidentified.

XX
PN WO9558947-A2.

XX
PD 18-NOV-1999.

XX
PF 12-MAY-1999; 99WO-US10361.

XX
PR 12-MAY-1998; 98US-0076404.

XX
PR 12-MAY-1998; 98US-0085092.

XX
PA (ISIS-) ISIS PHARM INC.

XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hostadler S, McNeil J;

XX
DR WPI: 2000-086439/07.

XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405bp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAUCUAGUUUACGAAAUAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 other;
XX
Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 4 gauncunnnuaagcccnangng 27
||:|:::|:||||| | | |
Db 7 gatctttt|gttaagccctagcg 30
XX
RESULT 10
AA71119
ID AA71119 standard; DNA; 42 BP.
XX
AC AA71119;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #125.
XX
Modulator: identification; molecular interaction; virtual library; ss.
XX
Unidentified.
XX
OS Unidentified.
XX
PN W0958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeill J;
XX
DR WPI: 2000-086439/07.
XX
Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405bp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAUCUAGUUUACGAAAUAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 T; 0 other;
XX
Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 4 gauncunnnuaagcccnangng 27
||:|:::|:||||| | | |
Db 7 gatctttt|gttaagccctagcg 30
XX
RESULT 11
AA71120
ID AA71120 standard; DNA; 42 BP.
XX
AC AA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
Modulator: identification; molecular interaction; virtual library; ss.
XX
Unidentified.
XX
OS Unidentified.
XX
PN W0958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeill J;
XX
DR WPI: 2000-086439/07.
XX
Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405bp; English.

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 3 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUACACAAAADC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

CC Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match

Best Local Similarity 62.1%; Score 18; DB 21; Length 42;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagccnangng 27
DB 7 gattcttttgaagcccaagg 30
||:|::|:||||| | | |
| | | | | | | | | |

RESULT 12

AAA71121
ID AAA71121 standard; DNA; 42 BP.

AC AAA71121;

DT 27-APR-2001 (first entry)

XX Molecular interaction site DNA #127.

XX Modulator; identification; molecular interaction; virtual library; ss.

OS Unidentified.

PN W09958947-A2.

PD 18-NOV-1999.

PF 12-MAY-1999; 99WO-US10361.

PR 12-MAY-1998; 98US-0076404.

PR 12-MAY-1998; 98US-0085092.

PA (ISIS-) ISIS PHARM INC.

PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hostadler S, McNeil J;

DR WPI; 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405bp; English.

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 3 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUACACAAAADC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

CC Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match

Best Local Similarity 62.1%; Score 18; DB 21; Length 42;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagccnangng 27
DB 7 gattcttttgaagcccaagg 30
||:|::|:||||| | | |
| | | | | | | | | |

RESULT 13

AAA71123
ID AAA71123 standard; DNA; 42 BP.

AC AAA71123;

DT 27-APR-2001 (first entry)

XX Molecular interaction site DNA #129.

XX Modulator; identification; molecular interaction; virtual library; ss.

OS Unidentified.

PN W09958947-A2.

PD 18-NOV-1999.

PF 12-MAY-1999; 99WO-US10361.

PR 12-MAY-1998; 98US-0076404.

PR 12-MAY-1998; 98US-0085092.

PA (ISIS-) ISIS PHARM INC.

PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hostadler S, McNeil J;

DR WPI; 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

Example 7; Figure 125; 405pp; English.

This invention describes a novel method for identifying compounds which modulate the activity of a target biomolecule. The method uses 3-dimensional representations of the biomolecule and a library of compounds and comprises (a) identifying at least one molecular interaction site of the target RNA; (b) generating in silico a virtual library of compounds predicted or calculated to interact with the molecular interaction site; and (c) comparing 3-dimensional (3-D) representations of the target RNA with members of the virtual library of compounds to generate a hierarchy of the compounds ranked in accordance with their respective ability to form physical interactions with the molecular interaction site. The method also describes (1) RNA comprising a joined sequence of at least 24 nucleotides but not more than 70 nucleotides and having secondary structure defined by: (a) 3 nucleotides forming a first side of a first double stranded (ds) region; (b) 2 nucleotides forming a first side of an internal loop region; (c) 4 nucleotides forming a first side of a second ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4 nucleotides forming a second side of the second ds region; (f) 4 nucleotides forming a second side of the internal loop region; and (g) 3 nucleotides forming a second side of the first ds region; (2) a purified and isolated RNA fragment comprising the human sequence UUUACAACUAUCUAGUUUACGAAAUAC (II). The methods and products can be used for identifying agents which modulate the activity of biomolecules, particularly RNA. Such agents can be used as pharmaceutical, agricultural or industrial compounds.

Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagccnang 27
||:|::|:||||| | | |
Db 7 gattcttttgtaagcctaagg 30

RESULT 14

AAA71124
ID AAA71124 standard; DNA; 42 BP.

XX AAA71124;

XX 27-APR-2001 (first entry)

XX Molecular interaction site DNA #130.

XX Modulator: identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX Hotstadler S, McNeill J;

XX WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
XX used to provide compounds which can be used as pharmacological,
XX agricultural and industrial compounds -
XX

Example 7; Figure 125; 405pp; English.

This invention describes a novel method for identifying compounds which modulate the activity of a target biomolecule. The method uses 3-dimensional representations of the biomolecule and a library of compounds and comprises (a) identifying at least one molecular interaction site of the target RNA; (b) generating in silico a virtual library of compounds predicted or calculated to interact with the molecular interaction site; and (c) comparing 3-dimensional (3-D) representations of the target RNA with members of the virtual library of compounds to generate a hierarchy of the compounds ranked in accordance with their respective ability to form physical interactions with the molecular interaction site. The method also describes (1) RNA comprising a joined sequence of at least 24 nucleotides but not more than 70 nucleotides and having secondary structure defined by: (a) 3 nucleotides forming a first side of a first double stranded (ds) region; (b) 2 nucleotides forming a first side of an internal loop region; (c) 4 nucleotides forming a first side of a second ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4 nucleotides forming a second side of the second ds region; (f) 4 nucleotides forming a second side of the internal loop region; and (g) 3 nucleotides forming a second side of the first ds region; (2) a purified and isolated RNA fragment comprising the human sequence UUUACAACUAUCUAGUUUACGAAAUAC (II). The methods and products can be used for identifying agents which modulate the activity of biomolecules, particularly RNA. Such agents can be used as pharmaceutical, agricultural or industrial compounds.

Sequence 42 BP; 11 A; 10 C; 7 G; 14 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuununguaagccnang 27
||:|::|:||||| | | |
Db 7 gattcttttgtaagcctaagg 30

RESULT 15

AAA71126
ID AAA71126 standard; RNA; 42 BP.

XX AAA71126;

XX 27-APR-2001 (first entry)

XX Molecular interaction site RNA #195.

XX Modulator: identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

XX Hotstadler S, McNeill J;

XX WPI: 2000-086439/07.

XX Identifying compounds which modulate activity of target biomolecules,
XX used to provide compounds which can be used as pharmacological,
XX agricultural and industrial compounds -
XX

PS Example 7; Figure 126; 405bp; English.

xx
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUCACACUAUUCUAGUUUCAGAAAAUUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

xx
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 U; 0 other;

Query Match

Best Local Similarity 62.1%; Score 18; DB 21; Length 42;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gauncuuuuuuuagcccnangng 27
||| |||| |||| |||| ||||
Db 7 gaucuuuuuuuagcccuacgag 30

Search completed: October 2, 2001, 05:03:56
Job Time: 3808 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:00:40 ; Search time 1315.38 seconds
(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-24

Perfect score: 29

Sequence: 1 uaugauuuuuuuuuaagccuagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 773874588 residues 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_cm:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hcgo_hum:*
20: em_hcgo_inv:*
21: em_hcgo_rod:*
22: em_hcgo_hum1:*
23: em_hcgo_hum2:*
24: em_hcgo_hum3:*
25: em_hcgo_hum4:*
26: em_hcgo_hum5:*
27: em_hcgo_hum6:*
28: em_hcgo_hum7:*
29: em_hcgo_hum8:*
30: em_hcgo_inv1:*
31: em_hcgo_inv2:*
32: em_hcgo_other:*
33: em_hcgo_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_cm:*
43: em_or:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	29	100.0	209 10	182323 Sequence 8
2	29	100.0	292 54	G06364 human STR W
3	29	100.0	689 9	A17717 Plasmid pTG
4	29	100.0	698 9	A02159 Synthetic g
5	29	100.0	698 9	A04681 Synthetic g
6	29	100.0	698 9	A14089 Synthetic D
7	29	100.0	722 10	I03416 Sequence 2
8	29	100.0	730 91	K03174 Ape (gibbon

9 29 100.0 750 10 101391 101391 Sequence 3
10 29 100.0 769 10 E01460 E01460 CDNA encodi
11 29 100.0 769 10 E02201 E02201 DNA encodi
12 29 100.0 782 10 101101 101101 Sequence 2
13 29 100.0 784 9 A06879 A06879 Artificial
14 29 100.0 784 9 A14095 A14095 Synthetic D
15 29 100.0 784 9 A17732 A17732 Plasmid prg
16 29 100.0 784 10 E00250 E00250 DNA coding
17 29 100.0 784 10 108366 108366 Sequence 2
18 29 100.0 788 10 107944 107944 Sequence 1
19 29 100.0 790 10 104498 104498 Sequence 1
20 29 100.0 790 10 108377 108377 Sequence 1
21 29 100.0 794 10 E00210 E00210 CDNA encodi
22 29 100.0 794 10 E00211 E00211 CDNA encodi
23 29 100.0 794 10 E02011 E02011 CDNA encodi
24 29 100.0 794 10 E02018 E02018 CDNA encodi
25 29 100.0 795 10 E00214 E00214 CDNA encodi
26 29 100.0 795 10 E00267 E00267 DNA sequenc
27 29 100.0 795 10 E00271 E00271 DNA sequenc
28 29 100.0 801 9 A04961 A04961 Artificial
29 29 100.0 801 9 A06759 A06759 Artificial
30 29 100.0 801 10 E00336 E00336 Human inter
31 29 100.0 801 10 101197 101197 Sequence 1
32 29 100.0 801 10 104318 104318 Sequence 3
33 29 100.0 801 93 H00564 H00564 Human mRNA
34 29 100.0 810 10 E00978 E00978 CDNA sequen
35 29 100.0 810 10 101387 101387 Sequence 1
36 29 100.0 812 10 E00216 E00216 DNA encodi
37 29 100.0 812 10 E02540 E02540 CDNA encodi
38 29 100.0 812 97 S77834 S77834 Homo saplen
39 29 100.0 825 97 HSU25676 HSU25676 Homo saplen
40 29 100.0 844 93 H01586 H01586 Human mRNA
41 29 100.0 1028 97 S82652 S82652 Interleukin
42 29 100.0 5561 97 X00695 X00695 Human inter
43 29 100.0 6684 93 AF359393 AF359393 Homo sapi
44 29 100.0 6752 89 AR031529 AR031529 Sequence
c 45 29 100.0 8491 9 AR031529 AR031529 Sequence

ALIGNMENTS

RESULT 1
182323
LOCUS 182323 209 bp DNA
DEFINITION Sequence 8 from patent US 5712126.
ACCESSION 182323
VERSION 182323.1 GI:3210620
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 209)
TITLES Weissman,S.M. and Prashar,Y.
JOURNAL Analysis of gene expression by display of 3-end restriction
FEATURES fragments of CDNA
PATENT: US 5712126-A 8 27 -JAN-1998;
source Location/Qualifiers
1..209
BASE COUNT 56 a 30 c 33 g 90 t
ORIGIN /organism="unknown"

Query Match 100.0%; Score 29; DB 10; Length 209;
Best Local Similarity 58.6%; Pred. No. 0.0038;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuuaagccuaggguu 29
:||||:||||:||||:||||:||||:||||:
Db 33 TATGATCTCTTTTGTAGACCCCTAGGGGCT 61

RESULT 2
G06364 292 bp DNA STS 19-OCT-1995
LOCUS G06364
DEFINITION human STS WI-7035.
ACCESSION G06364
VERSION G06364.1 GI:859609
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene collection.
ORGANISM Homo sapiens
Eukaryotes; mitochondria eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choeanata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivora; Homiidae; Homo.
REFERENCE 1 (bases 1 to 292)
AUTHORS Hudson,T.
TITLES Whitehead Institute/MIT Center for Genome Research; Physically
Mapped ESTs
JOURNAL Mapped ESTs
COMMENT Unpublished (1995)

Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: TAATTAAGTCCTCCACTTAAC
Primer B: ATTGTGGATTAATTAAGTAAACA
STS size: 200
PCR Profile:

Presoak:

Denaturation:

Annealing: 56 degrees C

Polymerization:

PCR Cycles: 35

Thermal Cycler:

Protocol:

Template: 10 ng

Primer: each 5 pm

dNTPs: each 4 mM

Taq Polymerase: 0.025 units/ul

Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCL: 10 mM

pH: 9.3

Prepared with primer pairs derived from V00564 -- Unigene.
location/Qualifiers

FEATURES

source

STS

primer_bind

primer_bind

BASE COUNT 92 a 32 c 32 g 124 t 12 others

ORIGIN

Query Match 100.0%; Score 29; DB 54; Length 292;
Best Local Similarity 58.6%; Pred. No. 0.0039;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuuaagccuaggguu 29
:||||:||||:||||:||||:||||:||||:
Db 141 TATGATCTCTTTTGTAGACCCCTAGGGGCT 169

RESULT 3
A17717

LOCUS A17717 689 bp mRNA PAT 17-MAY-1994
DEFINITION Plasmid pTG 26 mRNA for human IL-2.
ACCESSION A17717
VERSION A17717.1 GI:513949
KEYWORDS
SOURCE unidentified.
ORGANISM plasmid unidentified
REFERENCE 1 (bases 1 to 689)
AUTHORS Kieny,M.P., Sondermeyer,P. and Lecocq,J.P.
TITLE Expression of human IL-2 in mammalian cells by a recombinant pox virus
JOURNAL Patent: EP 0206939-A 1 30-DEC-1986;
TRANSGENE S.A
FEATURES
source Location/Qualifiers
1..689
/organism="unidentified"
/plasmid="pTG26"
/db_xref="taxon:32644"
1..384
/partial
/codon_start=1
/product="human interleukin 2"
/protein_id="CAA01347.1"
/db_xref="GI:513950"
/translation="TKTQLOLEHLLDLMQTLNGINNYKNPKLFRMLTFKPYMPKKA
TELKHLOCLEELKPLEEVNLAOSKNFHLRDRDISNINVIYELKSETFPMEYKA
DETATVEFLNRMTIFCOSIISTLR"
BASE COUNT 251 a 117 c 99 g 222 t
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 689;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 uaugauucuuuuuuaagccuagggcgcu 29
:||||:||||:||||:||||:||||:||||:
Db 525 TATGATTCCTTTTGTAAAGCCCTAGGGGCT 553

RESULT 4
LOCUS A02159 698 bp DNA PAT 21-MAY-1993
DEFINITION Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION A02159
VERSION A02159.1 GI:412310
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 698)
TITLE VECTOR FOR THE EXPRESSION IN YEASTS OF INTERLEUKINE-2, TRANSFORMED
JOURNAL YEASTS AND METHOD FOR PREPARING INTERLEUKINE-2
FEATURES Patent: WO 8503723-A 6 29-AUG-1985;
source Location/Qualifiers
1..698
/organism="synthetic construct"
/db_xref="taxon:32630"
10..393
/partial
/codon_start=1
/transl_table=11
/product="interleukin 2 (IL-2)"
/protein_id="CAA00227.1"
/db_xref="GI:412311"
/translation="TKKTQLOLEHLLDLMQTLNGINNYKNPKLFRMLTFKPYMPKKA
TELKHLOCLEELKPLEEVNLAOSKNFHLRDRDISNINVIYELKSETFPMEYKA
DETATVEFLNRMTIFCOSIISTLR"
BASE COUNT 251 a 117 c 107 g 222 t 1 others
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 698;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 uaugauucuuuuuuaagccuagggcgcu 29
:||||:||||:||||:||||:||||:||||:
Db 534 TATGATTCCTTTTGTAAAGCCCTAGGGGCT 562

RESULT 5
LOCUS A04681 698 bp DNA PAT 24-MAY-1993
DEFINITION Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION A04681
VERSION A04681.1 GI:412434
KEYWORDS
SOURCE interleukin 2.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 698)
AUTHORS Lemoine,Y., Sondermeyer,P., Loison,G., Aigle,M. and Lecocq,J.P.
TITLE Yeast-expression vectors for interleukin-2, transformed yeasts and
JOURNAL process for the preparation of interleukin-2
TRANSGENE S.A
FEATURES Patent: EP 0152358-A 5 21-AUG-1985;
source Location/Qualifiers
1..698
/organism="synthetic construct"
/db_xref="taxon:32630"
10..393
/partial
/codon_start=1
/transl_table=11
/product="interleukin 2 (IL-2)"
/protein_id="CAA00377.1"
/db_xref="GI:412435"
/translation="TKKTQLOLEHLLDLMQTLNGINNYKNPKLFRMLTFKPYMPKKA
TELKHLOCLEELKPLEEVNLAOSKNFHLRDRDISNINVIYELKSETFPMEYKA
DETATVEFLNRMTIFCOSIISTLR"
BASE COUNT 251 a 118 c 106 g 222 t 1 others
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 698;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 uaugauucuuuuuuaagccuagggcgcu 29
:||||:||||:||||:||||:||||:||||:
Db 534 TATGATTCCTTTTGTAAAGCCCTAGGGGCT 562

RESULT 6
LOCUS A14089 698 bp DNA PAT 28-FEB-1994
DEFINITION Synthetic DNA (pTG26) for human interleukin-2 (partial).
ACCESSION A14089
VERSION A14089.1 GI:491760
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 698)
TITLE EXPRESSION OF HUMAN IL-2 IN MAMMAL CELLS BY MEANS OF A RECOMBINED
JOURNAL POXVIRUS
FEATURES Patent: WO 8607610-A 1 31-DEC-1986;
source Location/Qualifiers
1..698
/organism="synthetic construct"
/db_xref="taxon:32630"
10..393
/partial
/codon_start=1
/transl_table=11
/product="interleukin 2 (IL-2)"
/protein_id="CAA00377.1"
/db_xref="GI:412435"
/translation="TKKTQLOLEHLLDLMQTLNGINNYKNPKLFRMLTFKPYMPKKA
TELKHLOCLEELKPLEEVNLAOSKNFHLRDRDISNINVIYELKSETFPMEYKA
DETATVEFLNRMTIFCOSIISTLR"
BASE COUNT 251 a 118 c 106 g 222 t 1 others
ORIGIN

```
/partial
/codon_start=1
/translation_start=1
/product="interleukin 2"
/protein_id="CAA01145.1"
/db_xref="GI:491761"
/translation="TKKIQLOLEHLIDLOMILNGINNYKPKITRLTKFYMPKKA
TELKHLQLEELKPLEVNLAKSKNPHLRPRDLISNINIVLELKGSETTCEYA
DETATVEFPLNRMTTFCOSIISTLT"
DEFINITION 251 a 117 c 108 g 222 t
ORIGIN
BASE COUNT 251 a 117 c 108 g 222 t

Query Match 100.0%; Score 29; DB 9; Length 698;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 naugauuuuuuuuagaccuaggggcu 29
:||||:||||:||||:||||:||||:
Db 534 TATGATTCCTTTTGTGAAGCCCTAGGGGCT 562

RESULT 7
LOCUS 103416 722 bp ss-DNA PAT 21-MAY-1993
DEFINITION Sequence 2 from Patent US 4882282.
ACCESSION 103416
VERSION 103416.1 GI:270625
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 722)
Anderson, D.M., Baker, P.E., Cantrell, M.A., Cerretti, D.P.,
Cosman, D.J., Gimpel, S.D., Grabstein, K.H., Larsen, A.D. and
McKereghan, K.N.
DNA sequences encoding bovine interleukin-2
Patent: US 4882282-A 2 21-NOV-1989;
Immunex Corporation;
Seattle, WA
FEATURES
source Location/Qualifiers
BASE COUNT 245 a 130 c 112 g 235 t
ORIGIN

Query Match 100.0%; Score 29; DB 10; Length 722;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 naugauuuuuuuuagaccuaggggcu 29
:||||:||||:||||:||||:||||:
Db 610 TATGATTCCTTTTGTGAAGCCCTAGGGGCT 638

RESULT 8
LOCUS 730 bp mRNA PRI 27-APR-1993
DEFINITION Ape (gibbon) interleukin 2 mRNA.
ACCESSION K03174
VERSION K03174.1 GI:177012
KEYWORDS interleukin; interleukin 2.
SOURCE Gibbon ('normal gibbon'), cDNA to mRNA, clone CM1.
ORGANISM Hyllobates lar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hylodactylidae; Hyllobates.
REFERENCE 1 (bases 1 to 730)
Chen, S.J., Holbrook, N.J., Mitchell, K.F., Vallone, C.A.,
Greenard, J.S., Crabtree, G.R. and Lin, Y.
A vital long terminal repeat in the interleukin 2 gene of a cell
line that constitutively produces interleukin 2
Proc. Natl. Acad. Sci. U.S.A. 82, 7284-7288 (1985)
JOURNAL
```

```
MEDLINE 86042650
COMMENT [1] also sequenced the IL-2 mRNA from gibbon cell line MIA 144 (see
separate entry).
FEATURES
source Location/Qualifiers
1..730
/organism="Hyllobates lar"
/db_xref="taxon:9580"
<1..>729
/feature="IL-2 mRNA"
25..486
/feature="interleukin-2"
/codon_start=1
/protein_id="AAA35453.1"
/db_xref="GI:177013"
/translation="MYRQLLSCLATSLATVNSAPYSSTKKTLQLOLEHLIDLOMI
LNGINNYKPKITRLTKFYMPKATLHLOLEELKPLEVNLAKSKNPHLR
RDLISNINIVLELKGSETTFCMEYADETATVEFPLNRMTTFCOSIISTLT"
variation 241
/note="c in CM1; t in CM9"
BASE COUNT 247 a 138 c 110 g 235 t
ORIGIN 180 bp upstream of HinfI site.

Query Match 100.0%; Score 29; DB 9; Length 730;
Best Local Similarity 58.6%; Pred. No. 0.0043;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 naugauuuuuuuuagaccuaggggcu 29
:||||:||||:||||:||||:||||:
Db 627 TATGATTCCTTTTGTGAAGCCCTAGGGGCT 655

RESULT 9
LOCUS 101391 750 bp ss-DNA PAT 21-MAY-1993
DEFINITION Sequence 3 from Patent US 4851512.
ACCESSION 101391
VERSION 101391.1 GI:270228
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 750)
Miyaji, H. and Itoh, S.
Novel human interleukin-2 polypeptide derivative
Patent: US 4851512-A 3 25-JUL-1989;
Kyowa Hakko Kogyo Co., Ltd.;
Tokyo, JP;
FEATURES
source Location/Qualifiers
BASE COUNT 276 a 131 c 104 g 239 t
ORIGIN

Query Match 100.0%; Score 29; DB 10; Length 750;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Oy 1 naugauuuuuuuuagaccuaggggcu 29
:||||:||||:||||:||||:||||:
Db 588 TATGATTCCTTTTGTGAAGCCCTAGGGGCT 616

RESULT 10
LOCUS 769 bp RNA PAT 29-SEP-1997
DEFINITION cDNA encoding human interleukin 2.
ACCESSION E01460
VERSION E01460.1 GI:2169716
KEYWORDS JP 1987265992-A/1.
SOURCE Homo sapiens.
```


ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Onomichi, K., Eto, Y. and Shibai, H.
TITLE PRODUCTION OF VALUABLE SUBSTANCE USING HUMAN CELL
JOURNAL Patent: JP 1987265992-A 1 18-NOV-1987;
AJINOMOTO CO INC
OS Human
COMMENT PN JP 1987265992-A/1
PD 18-NOV-1987
PE 12-MAY-1986 JP 1986107981
PI ONOMICHI KAZUYA, ETO YUZURU, SHIBAI HIROSHIRO PC
C12P21/00.C12N15/00.C12P21/02.(C12P21/02.C12R1.91); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=PSDI;
FH Key Location/Qualifiers
FT 5'UTR 1..2
FT CDS 3..507
FT 508..814 /product='human interleukin 2' FT 3'UTR
FT mat_peptide 3..504
FT Location/Qualifiers
source 1..769
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 280 a 132 c 112 g 245 t
ORIGIN

Query Match 100.0%; Score 29; DB 10; Length 769;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 naugaucuuuuuuaagcccuagggcu 29
Db 605 TATGATTCTTTTGTAAAGCCCTAGGGCT 633

RESULT 11
LOCUS E02201 769 bp RNA PAT 29-SEP-1997
DEFINITION DNA encoding interleukin2 (IL-2).
ACCESSION E02201
VERSION E02201.1 GI:2170439
KEYWORDS JP 1990009388-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Murata, M., Eto, Y. and Shibai, H.
TITLE PRODUCTION OF PHYSIOLOGICALLY ACTIVE PROTEIN
JOURNAL Patent: JP 1990009388-A 2 12-JAN-1990;
AJINOMOTO CO INC
OS Homo sapiens
COMMENT PN JP 1990009388-A/2
PD 12-JAN-1990
PE 08-JUL-1988 JP 1988170142
PR 09-MAR-1988 JP 88P 552270
PI MURATA MASAHIRO, ETO YUZURU, SHIBAI HIROSHIRO PC
C12P21/00.C12N15/12.C12N15/26.C12P21/02.(C12P21/02.C12R1.91); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: library=PSD(1)/IL-2;
CC Feature is identified by experimental;

FEATURES
source Location/Qualifiers
FT 3'UTR 465..769.
FT mat_peptide 3..461
FT /product='interleukin2'
FT /note='IL-2'
FT /note='IL-2'
FT Location/Qualifiers
source 1..769
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 280 a 132 c 112 g 245 t
ORIGIN

Query Match 100.0%; Score 29; DB 10; Length 769;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 naugaucuuuuuuaagcccuagggcu 29
Db 605 TATGATTCTTTTGTAAAGCCCTAGGGCT 633

RESULT 12
LOCUS I01101 782 bp ss-DNA PAT 21-MAY-1993
DEFINITION Sequence 2 from Patent US 4761375.
ACCESSION I01101
VERSION I01101.1 GI:269302
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 782)
AUTHORS Clark, S.C.
TITLE Human Interleukin-2 cDNA sequence
JOURNAL Patent: US 4761375-A 2 02-AUG-1988;
Genetics Institute, Inc.;
Cambridge, MA
FEATURES
source Location/Qualifiers
1..782
/organism='unknown'
BASE COUNT 272 a 143 c 113 g 253 t 1 others
ORIGIN

Query Match 100.0%; Score 29; DB 10; Length 782;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 naugaucuuuuuuaagcccuagggcu 29
Db 641 TATGATTCTTTTGTAAAGCCCTAGGGCT 669

RESULT 13
LOCUS A06879 784 bp DNA PAT 10-NOV-1993
DEFINITION Artificial sequence for interleukin-related protein.
ACCESSION A06879
VERSION A06879.1 GI:490471
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 784)
AUTHORS Fiers, W.C. and Devos, R.R.
TITLE DNA sequences, recombinant DNA molecules and processes for
producing human interleukin two-like polypeptides

JOURNAL Patent: EP 0118977-A 15 19-SEP-1984;
Biogen, Inc.; BIOGEN, INC
FEATURES Location/Qualifiers
source 1..784
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS 48..509
/codon_start=1
/translation="interleukin-related protein"
/product="interleukin-related protein"
/protein_id="CA00601.1"
/db_xref="GI:490472"
/translation="MYRQLSCIALSLAVTNSAPTSSTKKTQLOLEHLLDLOMI
LNGINNYKNPKLTMLTFKFKYMPKATELKHLOCLEELKPLEEVNLNLAOSKNFHLRP
RDLISNINIVLELKGSETFPMCEYAHETATIVFELNRMWTFCCSIISTLT"

BASE COUNT 267 a 146 c 115 g 256 t
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 784;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuagaccuagggccu 29
:||||:||||:||||:||||:||||:||||:
DB 650 TATGATCTCTTTGTAGCCCTAGGGGCT 678

RESULT 14
A14095 784 bp DNA PAT 28-FEB-1994
LOCUS
DEFINITION Synthetic DNA (PTG36) for human interleukin-2 (partial).
ACCESSION A14095
VERSION A14095.1 GI:491762
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 784)
AUTHORS
TITLE EXPRESSION OF HUMAN IL-2 IN MAMMAL CELLS BY MEANS OF A RECOMBINED
POXVIRUS
JOURNAL Patent: WO 8607610-A 7 31-DEC-1986;
FEATURES Location/Qualifiers
source 1..784
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS 13..474
/codon_start=1
/translation="interleukin 2"
/product="interleukin 2"
/protein_id="CA01146.1"
/db_xref="GI:491763"
/translation="MYRQLSCIALSLAVTNSAPTSSTKKTQLOLEHLLDLOMI
LNGINNYKNPKLTMLTFKFKYMPKATELKHLOCLEELKPLEEVNLNLAOSKNFHLRP
RDLISNINIVLELKGSETFPMCEYAHETATIVFELNRMWTFCCSIISTLT"

BASE COUNT 273 a 148 c 116 g 247 t
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 784;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuagaccuagggccu 29
:||||:||||:||||:||||:||||:||||:
DB 615 TATGATCTCTTTGTAGCCCTAGGGGCT 643

RESULT 15
A17732 784 bp mRNA PAT 17-MAY-1994
LOCUS
DEFINITION Plasmid PTG 36 mRNA for IL-2.

ACCESSION A17732
VERSION A17732.1 GI:513957
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 784)
AUTHORS Kieny, M.P., Sandermeier, P. and Lecocq, J.P.
TITLE Expression of human IL-2 in mammalian cells by a recombined pox
virus
JOURNAL Patent: EP 0206939-A 16 30-DEC-1986;
TRANSGENE S.A
FEATURES Location/Qualifiers
source 1..784
/organism="unidentified"
/db_xref="taxon:32644"
source 13..784
/organism="unidentified"
/plasmid="PTG36"
/db_xref="taxon:32644"
CDS 13..474
/codon_start=1
/product="interleukin-2"
/protein_id="CA01348.1"
/db_xref="GI:513958"
/translation="MYRQLSCIALSLAVTNSAPTSSTKKTQLOLEHLLDLOMI
LNGINNYKNPKLTMLTFKFKYMPKATELKHLOCLEELKPLEEVNLNLAOSKNFHLRP
RDLISNINIVLELKGSETFPMCEYAHETATIVFELNRMWTFCCSIISTLT"

BASE COUNT 273 a 148 c 116 g 247 t
ORIGIN

Query Match 100.0%; Score 29; DB 9; Length 784;
Best Local Similarity 58.6%; Pred. No. 0.0044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuagaccuagggccu 29
:||||:||||:||||:||||:||||:||||:
DB 615 TATGATCTCTTTGTAGCCCTAGGGGCT 643

Search completed: October 2, 2001, 05:00:41
Job time: 3:28 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:01:46 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-24
29
Perfect score: 1 unagaucuuuuuuaagccuagggcu 29
Sequence: 1 unagaucuuuuuuaagccuagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	29	100.0	209 1 US-08-510-032A-8	Sequence 8, Appl
2	29	100.0	209 3 US-08-688-514-8	Sequence 8, Appl
3	29	100.0	801 6 5314995-8	Patent No. 5314995
4	29	100.0	8491 2 US-08-757-439-1	Sequence 1, Appl
5	18.4	63.4	2885 2 US-08-791-849A-15	Sequence 15, Appl
6	17.8	61.4	4093 4 US-07-861-458C-5	Sequence 5, Appl
7	17	58.6	8920 4 US-08-446-855A-1	Sequence 1, Appl
8	17	58.6	8920 4 US-08-446-855A-1	Sequence 1, Appl
9	16.4	56.6	512 2 US-08-628-428-7	Sequence 4, Appl
10	16.4	56.6	512 2 US-08-628-428-7	Sequence 4, Appl
11	16.4	56.6	51259 3 US-08-781-891-209	Sequence 209, App
12	16.4	56.6	51259 3 US-08-781-891-209	Sequence 209, App
13	16.2	55.9	1846 3 US-09-173-581-15	Sequence 15, Appl
14	16	55.2	223 3 US-08-621-018B-48	Sequence 48, Appl
15	16	55.2	223 3 US-08-621-018B-48	Sequence 48, Appl
16	16	55.2	230 3 US-08-621-018B-46	Sequence 46, Appl
17	16	55.2	251 3 US-08-621-018B-42	Sequence 42, Appl
18	16	55.2	260 3 US-08-621-018B-43	Sequence 43, Appl
19	16	55.2	9493 2 US-08-639-857-23	Sequence 23, Appl
20	15.8	54.5	15202 3 US-08-922-635-21	Sequence 21, Appl
21	15.8	54.5	36519 3 US-08-923-137-2	Sequence 2, Appl
22	15.6	53.8	1166 1 US-08-599-252-95	Sequence 95, Appl
23	15.6	53.8	1166 5 PCT-US96-06352-95	Sequence 95, Appl
24	15.6	53.8	1166 5 PCT-US96-06352-95	Sequence 95, Appl
25	15.6	53.8	1801 1 US-08-391-000-41	Sequence 41, Appl
26	15.6	53.8	1801 1 US-08-741-931-41	Sequence 41, Appl
27	15.6	53.8	1897 1 US-08-245-688-1	Sequence 1, Appl

28	15.6	53.8	1897 1 US-08-245-688-3	Sequence 3, Appl
29	15.6	53.8	1897 1 US-08-245-688-5	Sequence 5, Appl
30	15.6	53.8	1897 1 US-08-245-688-7	Sequence 7, Appl
31	15.6	53.8	1897 1 US-08-245-688-9	Sequence 9, Appl
32	15.6	53.8	1897 1 US-08-245-688-11	Sequence 11, Appl
33	15.6	53.8	2300 1 US-08-103-998-3	Sequence 3, Appl
34	15.6	53.8	2549 1 US-08-470-720-2	Sequence 2, Appl
35	15.6	53.8	3611 2 US-08-727-118-1	Sequence 1, Appl
36	15.6	53.8	4964 1 US-08-470-720-5	Sequence 5, Appl
37	15.6	53.8	7208 3 US-09-166-186-107	Sequence 107, App
38	15.6	53.8	7208 4 US-09-313-932-107	Sequence 107, App
39	15.4	53.1	107 1 US-08-441-591-45	Sequence 45, Appl
40	15.4	53.1	107 1 US-08-303-362A-45	Sequence 45, Appl
41	15.4	53.1	107 5 PCT-US95-05600-62	Sequence 62, Appl
42	15.4	53.1	246 3 US-08-617-860B-14	Sequence 14, Appl
43	15.4	53.1	340 3 US-08-441-971-13	Sequence 13, Appl
44	15.4	53.1	340 4 US-08-221-653-13	Sequence 13, Appl
45	15.4	53.1	340 4 US-08-442-144A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-510-032A-8
Sequence 8, Application US/08510032A
Patent No. 5712126
GENERAL INFORMATION:
APPLICANT: Sherman Weissman and Yarlindra Prashar
TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,032A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: Yale
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-510-032A-8

Query Match 100.0%; Score 29; DB 1; Length 209;
Best Local Similarity 58.6%; Pred. No. 7.4e-05;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 unagaucuuuuuuaagccuagggcu 29
DB 33 TATGATTCTTTTGTAAAGCCCTAGGGGCT 61

```
RESULT 2
US-08-688-514-8
; Sequence 8, Application US/08688514
; Patent No. 6010850
; GENERAL INFORMATION:
; APPLICANT: Sherman Weissman and Yairindra Prashar
; TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
; TITLE OF INVENTION: end Restriction Fragments of cDNA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,514
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: Yale
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-688-514-8

Query Match 100.0%; Score 29; DB 3; Length 209;
Best Local Similarity 58.6%; Pred. NO. 7.4e-05;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 uauguuncuuuuuagccuaggggcu 29
:||||:||||:||||:||||:||||:||||:
33 TATGATCTTTTGTAGCCCTAGGGGCT 61

RESULT 3
5314995-8
; Patent No. 5314995
; APPLICANT: FELL, HENRY P.; GAYLE, MARCIT A.
; TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
; BASED FUSION PROTEINS
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/468,390
; FILING DATE: 22-JAN-1990
; SEQ ID NO: 8
; LENGTH: 801
5314995-8

Query Match 100.0%; Score 29; DB 6; Length 801;
Best Local Similarity 58.6%; Pred. NO. 9.3e-05;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 uauguuncuuuuuagccuaggggcu 29
:||||:||||:||||:||||:||||:||||:
33 TATGATCTTTTGTAGCCCTAGGGGCT 61
```

```
Db 650 tatgattctttgtgaagccctaggggct 678

RESULT 4
US-08-757-439-1/c
; Sequence 1, Application US/08757439
; Patent No. 5866371
; GENERAL INFORMATION:
; APPLICANT: BADZLONG, Werner
; APPLICANT: BABERMANN, Paul
; APPLICANT: MOELLER, Joerg
; APPLICANT: ARETZ, Werner
; TITLE OF INVENTION: PROCESS FOR USING THE YEAST ADH II
; TITLE OF INVENTION: PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,439
; FILING DATE: 27-NOV-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19544233.4
; FILING DATE: 28-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 18748/303/HOCE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8491 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-757-439-1

Query Match 100.0%; Score 29; DB 2; Length 8491;
Best Local Similarity 58.6%; Pred. NO. 0.00014;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 uauguuncuuuuuagccuaggggcu 29
:||||:||||:||||:||||:||||:||||:
Db 6654 TATGATCTTTTGTAGCCCTAGGGGCT 6626

RESULT 5
US-08-791-849A-15/c
; Sequence 15, Application US/08791849A
; Patent No. 5914449
; GENERAL INFORMATION:
; APPLICANT: MAKOTO MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
```

```

      FILING DATE:
      ATTORNEY/AGENT INFORMATION:
        NAME: Clark, Paul T.
        REGISTRATION NUMBER: 30,162
        REFERENCE/DOCKET NUMBER: 04585/014001
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (617) 542-5070
          TELEFAX: (617) 542-8906
          TELEX: 200154
        INFORMATION FOR SEQ ID NO: 5:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 4093
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
US-07-861-458C-5

Query Match              61.4%; Score 17.8; DB 4; Length 4093;
Best Local Similarity    41.4%; Pred. No. 18;
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

OY   1 uauagucuuuuuuuaagccuaggcgcu 29
      :|||::|:::|||||::|||:
Db   1146 TATGATTCTTTTAAAGTCGTAGTGCT 1118

RESULT       7
US-08-446-855A-1
; Sequence 1, Application US/08446855A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria V
; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
; TITLE OF INVENTION: phosphate synthetase II
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,855A
; FILING DATE: 06-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 47-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic
; US-08-446-855A-1

Query Match              58.6%; Score 17; DB 2; Length 8920;
Best Local Similarity    40.0%; Pred. No. 47;

```


APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg, Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 56.6%; Score 16.4; DB 3; Length 51259;
Best Local Similarity 42.3%; Pred. No. 1.2e+02;
Matches 11; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 4 gaucuuuuuuuagcccuaggagcu 29

Db 26016 GTTCTTTTGGAGCAGCTAGAGCT 26041

RESULT 12
US-08-781-891-209/c

Sequence 209, Application US/08781891
Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Yang-Hui
APPLICANT: Yu, Chang-Eh
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg, Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 56.6%; Score 16.4; DB 3; Length 51259;
Best Local Similarity 42.3%; Pred. No. 1.2e+02;
Matches 11; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 4 gaucuuuuuuuagcccuaggagcu 29

Db 10495 GTTCTTTTGGAGCAGCTAGAGCT 10470

RESULT 13
US-09-173-581-15/c

Sequence 15, Application US/09173581A
Patent No. 6013455

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzai, Yalda
APPLICANT: Lu, Aina
TITLE OF INVENTION: Protein Kinase Homologs
FILE REFERENCE: PF-0614 US
CURRENT APPLICATION NUMBER: US/09/173,581A
CURRENT FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 15
LENGTH: 1846
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1567782
US-09-173-581-15

Query Match 55.9%; Score 16.2; DB 3; Length 1846;
Best Local Similarity 38.1%; Pred. No. 85;
Matches 8; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

QY 2 augauuuuuuuuagcccu 22

Db 155 ATGATCTTTTCTAAGGCA 135

RESULT 14

US-08-621-018B-48

Sequence 48, Application US/08621018B
Patent No. 6060239

GENERAL INFORMATION:

APPLICANT: Stuart, Susan G.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Murry, Lynn E.

1 TITLE OF INVENTION: CELLULREVIN HOMOLOGS
 2 NUMBER OF SEQUENCES: 51
 3 CORRESPONDENCE ADDRESS:
 4 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 5 STREET: 3174 Porter Drive
 6 CITY: Palo Alto
 7 STATE: CA
 8 COUNTRY: U.S.
 9 ZIP: 94304
 10
 11 COMPUTER READABLE FORM:
 12 MEDIUM TYPE: Diskette
 13 COMPUTER: IBM Compatible
 14 OPERATING SYSTEM: DOS
 15 SOFTWARE: FastSeq Version 1.5
 16 CURRENT APPLICATION DATA:
 17 APPLICATION NUMBER: US/08/621,018B
 18 FILING DATE: March 22, 1996
 19
 20 PRIOR APPLICATION DATA:
 21 APPLICATION NUMBER: 08/409,373
 22 FILING DATE: March 23, 1995
 23 ATTORNEY/AGENT INFORMATION:
 24 NAME: CERRONE, MICHAEL C.
 25 REGISTRATION NUMBER: 39,132
 26 REFERENCE/DOCKET NUMBER: PF-0029-1 CIP
 27 TELECOMMUNICATION INFORMATION:
 28 TELEPHONE: 650-855-0555
 29 TELEFAX: 650-845-4166
 30 INFORMATION FOR SEQ ID NO: 48:
 31 SEQUENCE CHARACTERISTICS:
 32 LENGTH: 223 base pairs
 33 TYPE: nucleic acid
 34 STRANDEDNESS: single
 35 TOPOLOGY: linear
 36 MOLECULE TYPE: cDNA
 37 IMMEDIATE SOURCE:
 38 LIBRARY: LUNGNOTO2
 39 CLONE: 375070
 40
 41 US-08-621-018B-48

```

Query Match          95.2%; Score 16; DB 3; Length 223;
Best Local Similarity 41.7%; Pred. No. 73;
Matches      10; Conservative    9; Mismatches      5; Indels      0; Gaps      0;

OY      6 uucuuuuuuaagcccaaggggcu 29
       : |::|::||| | | | |
Db      122 TACTTTTGTAAAGCACTACTGACT 145

SUBMIT 15
S-08-621-018B-49
Sequence 49, Application US/08621018B
Patent No. 6060239
GENERAL INFORMATION:
Applicant: Stuart, Susan G.
Applicant: Hawkins, Phillip R.
Applicant: Selhammer, Jeffrey J.
Applicant: Murty, Lynn E.
TITLE OF INVENTION: CELLULREVIN HOMOLOGS
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
```

```

1  APPLICATION NUMBER:  US/08/621,018B
2  FILING DATE:  March 22, 1996
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER:  08/409,373
5  FILING DATE:  March 23, 1995
6  ATTORNEY/AGENT INFORMATION:
7  NAME:  CERRONE, MICHAEL C.
8  REGISTRATION NUMBER:  39,132
9  REFERENCE/DOCKET NUMBER:  PF-0029-1 CIP
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE:  650-855-0555
12 TELEFAX:  650-845-4166
13 INFORMATION FOR SEQ. ID NO.:  49:
14 SEQUENCE CHARACTERISTICS:
15     LENGTH:  225 base pairs
16     TYPE:  nucleic acid
17     STRANDEDNESS:  single
18     TOPOLOGY:  linear
19     MOLECULE TYPE:  cDNA
20 IMMEDIATE SOURCE:
21     LIBRARY:  LATRNOT01
22     CLONE:  465647
23  US-08-018B-49

```

OY	6	uuuuuuuuuaagcccuaggaggu	29
	:	:	:
Db	70	TACTTTTGTGATGCCTACTACT	93

```
Search completed: October 2, 2001, 05:01:49
Job time: 5316 sec
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:38:38 ; Search time 1643.83 Seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-24

Perfect score: 29

Sequence: 1 uaugauucuuuuuugaagccuaggggcu 29

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Matches	13, Conservative	10, Mismatches	4, Indels	Gaps	0
QY	3	ugaauuuuuuuuuaagccuagggcu	29		
		: : : : : : : : : : : : : : : : :			
Db	394	TGACTTTTGTGTAAGCCCTGCGGCT	368		

[illegible]

PERENCE	1 (bases 1 to 494)
AUTHORS	Morgan, R.
TITLE	Chicken T cell ESTs
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robin Morgan

Townsend Hall, Newark, DE 19117, USA
Tel: 302-831-1341
Fax: 302-831-2822
Email: morganudel.edu
Clones can be ordered online at <http://www.chickent.udel.edu>.

```

source
1. 494
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="ptlc.pk001.120"
/clone_id="Chicken T cell cDNA library"
/sex="Male and Female"
/cell_type="splenic T cell"
/lab_host="E.coli DH10B"
/note="Vector: PB42AD"

BASE COUNT      151 a      82 c      96 g      143 t      22 others
ORIGIN

```

Query Match	67.6%	Score 19.6:	DB 113:	Length 494:
Best Local Similarity	46.4%	Pred. No. 2.1e+02:		
Matches 13: Conservative	9:	Mismatches 6:	Indels 0:	Gaps
2 augauuuuuuuuaagcccuagggcuc	29			
1:111: :::~: 1:1:				
Db 16 ATGAGCGCTTTTGTAAGCCCNACTACT	43			

RESULT	6
LOCUS	AZ349700
DEFINITION	AZ349700 619 bp DNA
ACCESSION	U0086K17
VERSION	Mouse 10kb plasmid U0086K17 library Mus musculus genomic clone U0086K17 R, DNA sequence.
KEYWORDS	AZ349700.1 GI:10428937
SOURCE	GSS.
ORGANISM	house mouse.
REFERENCE	Mus musculus.
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 619)
TITLE	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tlingay,A., von Niederhausern,A. and Wright,D. Weiss,R.
	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah, General Counsel

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0086 row: K column: 17
 Seq primer: CACACAGCAAAACACCTATGACC
 Class: plasmid ends
 High quality sequence stop: 619.
 location/qualifiers

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U086K17"
 /clone_id="Mouse 10kb plasmid u086k17"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
 /note="Vector: pMD42ny; Purified genomic DNA from M.
 Laboratory Mouse C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 (gii14732114(gb1AF129072.1)), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

	Query Match	67.6%	Score 19.6	DB 242	Length 619
	Best Local Similarity	53.8%	Pred. No. 2.1e+02		
	Matches 14	Conservative	8	Mismatches 4	Indels 0
				Gaps 0	
QY	4	gaauuuuuuuuuuagaccuagggguc	29		
db	289	gatttctttctgcgagaccctagtgct	314		

RESULT	7
A2899124	
LOCUS	
DEFINITION	A2899124 741 bp DNA
ACCESSION	RPCT-24-211B11.vv RPCT-24 Mus musculus genomic clone RPCT-24-211B11
VERSION	, DNA sequence.
KEYWORDS	A2899124
SOURCE	A2899124..1 GI:13218069
ORGANISM	GSS.
	house mouse.
	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 741)
AUTHORS	Zhao,S., Niemman,W., Malek,J., Shatman,S., Akinneti,B., Levins,M.,
	Tessagel,G., Geer,K., Kiol,M., Shvartsbeyn,A., Gebregeorgis,E.,
	Russell,D., de Jong,P. and Fraser,C.M.
TITLE	Mouse BAC End Sequences from Library RPCT-24.
JOURNAL	Unpublished (1999)


```
/organism="Mus musculus"  
/strain="C57BL/6J"
```

```

/db_xref="taxon:10090"
/clone="UUCGCM0367H06"
/clone_lib="Mouse 10kb plasmid UUCGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g11473211419b1A129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

	Query Match	Score	DB	Length
Best Local Similarity	48.3%	Pred. No. 2.6e+02		
Matches	14	Conservative	9	Mismatches 6; Indels 0; Gaps 0;
Qy	1	naugaucuuuuuuaagccuagggc	29	
Db	52	TACCAATCTTCTGCTGGCCCTTAGGGGCT	80	

RESULT	12
LOCUS	A0239971/c
DEFINITION	A0239971 309 bp DNA GSS 30-SEP-1998 CTR-HSP-2386F6.TR.1 CTR-HSP Homo sapiens genomic clone 2386F6, DNA
ACCESSION	A0239971
VERSION	A0239971
KEYWORDS	GI:3672169
SOURCE	GSS.
	human.

REFERENCE
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.
1 (bases 1 to 309)
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi.

TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL	Unpublished (1998)
COMMENT	Other_GSSS: CIT-HSP-2386F6.TF.1

Email: mdadams@tigr.org
 Clones are available from Research Genetics (inforesgen.com). BAC
 end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .309
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

[illegible]

RESULT	13			
LOCUS	AA072238/c			
DEFINITION	AA072238	348 bp	DNA	05-AUG-1998
ACCESSION	AA072238			
VERSION	AA072238			
KEYWORDS	AA072238.1	GI:391087		
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE
AUTHORS

1 (pages 1 to 348)
Maharaj, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.

TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE	99380589
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L

401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel.: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3024 row: L column: 3
Class: BAC ends
High quality sequence stop: 348.

FEATURES	Location/Qualifiers
SOURCE	1..348
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="Plate3024 Col=3 Row=L"
	/clone_lib="CTR Approved Human Genomic Sperm Library D"
	/sex="male"
	/note="Organ: sperm; Vector: pBelBac11; BAC clones in E-Coli DH10b"
BASE COUNT	111 a 59 c 49 g 129 t .
ORIGIN	

[illegible]

RESULT 14

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:03:56 ; Search time 122.48 Seconds
(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844B-24
Perfect score: 29
Sequence: 1 uaugauuuuuuuuagccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
N.Geneseq_0601:*

- 1: /SIDSI/gcgdata/geneseq/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	29	AAA70828	Molecular interact
2	29	100.0	42	AAA71123	Molecular interact
3	29	100.0	42	AAA71131	Molecular interact
4	29	100.0	209	AA761453	IL-2 CDNA 3'end.
5	29	100.0	209	AA259456	Interleukin-2 frag
6	29	100.0	698	AA550132	DNA sequence encod
7	29	100.0	698	AA560101	Sequence of human
8	29	100.0	722	AA701044	Sequence encoding
9	29	100.0	769	AA001764	Human interleukin-
10	29	100.0	784	AA40046	Sequence encoding
11	29	100.0	784	AA60787	Cloned sequence en

12	29	100.0	784	7	AA60102	Sequence of human
13	29	100.0	788	10	AA90253	Interleukin-2. Ho
14	29	100.0	794	5	AA40441	Gene encoding poly
15	29	100.0	794	10	AA90381	Recombinant human
16	29	100.0	800	6	AA50219	Sequence encoding
17	29	100.0	801	4	AA30031	Sequence of Interl
18	29	100.0	801	5	AA40057	Cloned human inter
19	29	100.0	801	5	AA40254	Sequence encoding
20	29	100.0	802	6	AA50279	DNA sequence conta
21	29	100.0	810	7	AA60840	Sequence encoding
22	29	100.0	844	21	AA620965	Human low adenosin
23	29	100.0	844	21	AA34843	Human adenosine re
24	29	100.0	5561	21	AA620964	Human low adenosin
25	29	100.0	5561	21	AA34842	Human adenosine re
26	29	100.0	22421	21	AA620966	Human low adenosin
27	29	100.0	22421	21	AA34844	Human adenosine re
28	29	100.0	45	21	AA70824	Molecular interact
29	29	96.6	46	21	AA71087	Molecular interact
30	28	96.6	46	21	AA71096	Molecular interact
31	28	96.6	46	21	AA71099	Molecular interact
32	28	96.6	46	21	AA71100	Molecular interact
33	28	96.6	46	21	AA71104	Molecular interact
34	25.8	89.0	42	21	AA71113	Molecular interact
35	25.8	89.0	42	21	AA71118	Molecular interact
36	25.8	89.0	42	21	AA71126	Molecular interact
37	24.8	85.5	46	21	AA71085	Molecular interact
38	24.8	85.5	46	21	AA71103	Molecular interact
39	23.8	82.1	42	21	AA71114	Molecular interact
40	23.8	82.1	42	21	AA71119	Molecular interact
41	23.8	82.1	42	21	AA71127	Molecular interact
42	23.8	82.1	46	21	AA71094	Molecular interact
43	23.8	82.1	46	21	AA71110	Molecular interact
44	23.2	80.0	29	21	AA70829	Molecular interact
45	23.2	80.0	29	21	AA70830	Molecular interact

ALIGNMENTS

RESULT	1
ID	AAA70828 standard; RNA: 29 BP.
XX	XX
AC	AAA70828;
XX	XX
DT	27-APR-2001 (first entry)
XX	XX
DE	Molecular interaction site RNA #28.
XX	XX
KW	Modulator; identification; molecular interaction; virtual library; ss.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO958947-A2.
XX	XX
PD	18-NOV-1999.
XX	XX
PF	12-MAY-1999; 99WO-US10361.
XX	XX
PR	12-MAY-1998; 98US-0076404.
XX	XX
PR	12-MAY-1998; 98US-0085092.
XX	XX
PA	(ISIS-) ISIS PHARM INC.
XX	XX
PI	Baker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX	XX
PI	Hofstadler S, McNeil J;
XX	XX
DR	WPI; 2000-086439/07.
XX	XX
PT	Identifying compounds which modulate activity of target biomolecules,
XX	XX
PT	used to provide compounds which can be used as pharmacological,
XX	XX
PT	agricultural and industrial compounds -
XX	XX

PS Claim 235; Page 235; 405pp; English

PS Claim 235; Page 235; 405pp; English.

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUUACACAAUUAUUCUUAUACACAAAUAUC (11'). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

CC Sequence 29 BP; 5 A; 5 C; 7 G; 12 U; 0 other;

Query Match	100.0%	Score 29	DB 21	Length 29
Best Local Similarity	100.0%	Pred. NC	0.00034	
Matches 29	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	naugauucuuuuuuguaagcccuaggggc	29
Db	1	naugauucuuuuuuguaagcccuaggggc	29

```

RESULT      2
AAA71123
ID   AAA71123 standard; DNA; 42 BP

```

AC AAA71123

DT 27-APR-2001 (first entry)

Molecular interaction site DNA #129.

Modulator; identification; molecular interaction; virtual library; ss.

OS Unidentified.

PN W09958947-A2

PD 18-NOV-1999.

PF 12-MAY-1999;

PR 12-MAY-1998; 98US-0076404.

XX

XX

PI Hofstadler S, McNeil J

DR WPI; 2000-086439/07.

PT Identifying compounds w/

PT agricultural and ind

PS Example 7; Figure 125; 405pp; English.

PS Example 7; Figure 125; 405bp; English.

XX

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCUGUUUACGAAUAUUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX

SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 58.6%; Pred. No. 0 00035;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 uaugauucuuuuuguaagcccuaggcgcu    25
        |||::||:::||::|||::|||::|||::
Db      4 tatgatcttltgtgaagccctaaggcgt    32
```

RESULT	3
AAA71131	
ID	AAA71131 standard; RNA; 42 BP

AC AAA71131;

DT 27-APR-2001 (first entry)

DE Molecular interaction site RNA #200.

KW Modulator; identification; molecular interaction; virtual library; ss.

OS Unidentified.

PN W09958947-A2.

PD 18-NOV-1999.

PF 12-MAY-1999;

PR 12-MAY-1998; 98US-0076404.

XX

XX

PI Hofstadler S, McNeill J

DR WPI; 2000-086439/07.

PT Identifying compounds w/

PT agricultural and industrial c

PS Example 1; Page 13; 46pp; English.

Db 33 tatgattcttttgaagccctaggagct 61

RESULT 6

AA50132 ID AAN50132 standard; DNA; 698 BP.

AC AAN50132;

DT 15-OCT-1991 (first entry)

DE DNA sequence encoding Interleukin-2.

KW Interleukin-2; ss.

OS Homo sapiens.

Key Location/Qualifiers
CDS 10..391
/tag= a
/label= Interleukin-2

EP152358-A.

PD 21-AUG-1985.

XX 16-FEB-1984; 84EP-0002350.

PR 16-FEB-1984; 84FR-0002350.

XX (TRAN-) TRANSGENE SA.

PI Lemolne Y, Sondermeyer P, Loison G, Aigle M, Lecocq JP;

XX WPI; 1985-204923/34.

DR P-PSDB; AAP50111.

XX Vectors for expression of Interleukin-2 yeasts - contg. yeast

PT promoter and sequence coding for Interleukin.

XX Disclosure; Fig 2; 18pp; French.

CC The human interleukin-2 gene is cloned and expressed in

CC Saccharomyces cerevisiae in a mature form without post-translational

CC modification.

XX Sequence 698 BP; 251 A; 119 C; 106 G; 222 T; 0 other;

Query Match 100.0%; Score 29; DB 6; Length 698;
Best Local Similarity 58.6%; Pred. No. 0.00052;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 naugauuucuuuuaagcccuaggggcu 29

Db 534 tatgattcttttgaagccctaggagct 562

RESULT 7

AA60101 ID AAN60101 standard; cDNA; 698 BP.

AC AAN60101;

DT 02-JUL-1991 (first entry)

DE Sequence of human Interleukin 2 (IL2) gene on plasmid pTG 26.

XX Pox virus vector; vaccinia; ss.

OS Homo sapiens.

Key Location/Qualifiers
CDS 10..393
/tag= a

FT conflict (T_45,G_45)

FT /tag= b

FT conflict (G_387,T_387)

FT /tag= c

FT /citation= 1

EP206939-A.

PD 30-DEC-1986.

XX 19-JUN-1986; 86EP-0401349.

PR 21-JUN-1985; 85FR-0009480.

XX (TRAN-) TRANSGENE SA.

PI Kieny MP, Sondermeyer P, Lecocq JP;

XX WPI; 1986-341222/52.

DR P-PSDB; AAP60088.

XX Genetically modified pox viruses - contg. DNA coding for human

XX Interleukin 2

XX Example; Fig 1; 33pp; French.

CC cDNA coding for IL2 in plasmid pTG26 is restructured by inserting a

CC 92bp synthetic peptide signal fragment, giving a plasmid pTG36. A

CC plasmid pTG186-POLY is prep'd. as in EP 206920. This is digested with

CC PstI and ligated with PstI-digested pTG36. The recombinants are used

CC to transform E. coli and selected to give plasmid pVIT2 (pTG188)

CC which is cloned in vaccinia virus. Citation 1 - Taniguchi, T. et al.

CC (1983) Nature, 302, p305. The conflicts are silent wrt translation.

XX Sequence 698 BP; 251 A; 117 C; 108 G; 222 T; 0 other;

Query Match 100.0%; Score 29; DB 7; Length 698;
Best Local Similarity 58.6%; Pred. No. 0.00052;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 naugauuucuuuuaagcccuaggggcu 29

Db 534 tatgattcttttgaagccctaggagct 562

RESULT 8

AA70144 ID AAN70144 standard; cDNA; 722 BP.

AC AAN70144;

DT 17-APR-1991 (first entry)

DE Sequence encoding human Interleukin-2 (IL-2).

XX Mastitis therapy; parasitic infection; lymphokine; ss.

OS Homo sapiens.

Key Location/Qualifiers
CDS 8..67
/tag= a

FT mat_peptide 68..469

FT misc_feature 11..719

FT /tag= c

FT /note= "used to probe bovine cDNA library"

```

XX EP215576-A.
PN
XX
XX 25-MAR-1987.
PD
XX
XX 15-AUG-1986; 86EP-0306303.
PF
XX
XX 31-JUL-1986; 86US-0888994.
PR
XX 16-AUG-1985; 85US-0766643.
PA
XX (IMMUNEX) IMMUNEX CORP.
XX
XX Anderson DM, Baker PE, Cantrell MA, Cerretti DP, Cosman DJ,
PI Gimpe SD, Grabstein KH, Larsen AD, McKereghan KN;
XX
XX WPI; 1987-081523/12.
DR P-PSDB; AAP70090.
XX
XX Recombinant bovine interleukin-2 - for treating mastitis,
XX respiratory and gastro-intestinal syndromes and parasitic
XX infections
XX
XX Disclosure; Fig 1; 47pp; English.
PS
XX
XX A bovine cDNA library was screened using a probe which was essentially
XX the entire length of human IL-2 cDNA (AAN70144). A single positive
XX host colony was identified. Plasmid bIL-2-4 was prepd. and the
XX nucleotide sequence detd. (AAN70145). The specific activity of the
XX bIL-2 is 4.5 x 10(4) units/ug protein.
XX
XX Sequence 722 BP; 245 A; 130 C; 112 G; 235 T; 0 other;
SQ

```

Query Match 100.0%; Score 29; DB 8; Length 722;
 Best Local Similarity 58.6%; Pred. No. 0.00052;
 Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 uaugauuuuuuuuagaccuagggcu 29
DB 610 tatgatcttttggtaagccctagggcct 638

```

RESULT 9
 ID AAO01764
 XX AAO01764 standard; DNA; 769 BP.
 AC AAO01764;
 XX
 XX 27-JUL-1990 (first entry)
 XX
 XX Human interleukin-2 gene.
 DE
 XX IL-2; dhfr; dihydrofolic acid reductase; differentiation.
 KM
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH CDS 3..506
 FT /*tag= a
 FT
 FT
 PN JP02009388-A.
 PD
 PD 12-JAN-1990.
 PD
 PF 08-JUL-1988; 88JP-0170142.
 PR
 PR 09-MAR-1988; 88JP-0055270.
 XX
 XX (AJIN) AJINOMOTO KK.
 PA
 PA WPI; 1990-055348/08.
 DR P-PSDB; AAR05414.
 XX

```

PT Physiologically active protein prepn.
PT by transforming plasmid having gene coding physiologically
PT active protein and gene of dihydrofolic acid reductase to hamster
PT ovary etc.
XX
XX Example 2; Fig 4; 12pp; Japanese.
XX
XX Gene may be expressed by transforming a dhfr negative strain of CHO cells
CC with an active IL-2 gene and dhfr carrying vector.
CC
XX Sequence 769 BP; 280 A; 131 C; 113 G; 245 T; 0 other;
SQ

```

Query Match 100.0%; Score 29; DB 11; Length 769;
 Best Local Similarity 58.6%; Pred. No. 0.00053;
 Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 uaugauuuuuuuuagaccuagggcu 29
DB 605 tatgatcttttggtaagccctagggcct 633

```

RESULT 10
 ID AAN40046
 XX AAN40046 standard; cDNA; 784 BP.
 AC AAN40046;
 XX
 XX 14-JAN-1992 (first entry)
 DT
 XX
 XX Sequence encoding interleukin-2 (IL-2) related polypeptide in
 DE clone pSV-hIL-2L-O.
 XX
 XX Diagnosis; therapy; cancer; tumour-specific cytotoxic cell; AIDS;
 KM multiple sclerosis; lupus; rheumatoid arthritis; herpes;
 KM viral disease; lymphokine; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH CDS 48..509
 FT /*tag= a
 FT polyA_signal 771..776
 FT /*tag= b
 XX
 XX EP118977-A.
 PN
 XX
 XX 19-SEP-1984.
 PD
 PD 25-JAN-1984; 84EP-0300439.
 PF
 XX 10-JUN-1983; 83GB-0015981.
 PR 08-FEB-1983; 83GB-0003383.
 XX
 XX (BIOJ) BIOGEN NV.
 PA
 XX
 XX Fiers WC, Devos RR;
 PI
 XX WPI; 1984-232548/38.
 DR P-PSDB; AAP40055.
 DR
 XX
 XX Prodn. of human interleukin 2-like polypeptide(s) - useful
 PT instead of IL-2 for stimulating the immune system etc.
 PT
 XX
 XX Disclosure; Fig 4; 69pp; English.
 PS
 XX
 XX The DNA sequence is esp. selected from a human chromosomal gene bank,
 CC e.g. it is a hIL-2 related portion of lambda CHA-ghIL-2-1 or -2, or
 CC of lambda I47-ghIL-2-1, -2 or -3. Transformed hosts are also claimed,
 CC esp. E.coli, P.spp.; B.subtilis, B.steartotherophilus. IL-2-like
 CC polypeptides are also claimed. AAN40046 is an IL-2 related DNA sq of
 CC the invention. It is missing at least part of the coding region for
 CC the amino acids of the putative signal sequence of pre hIL-2 and the

DT 01-NOV-1989 (first entry)
XX Interleukin-2.
DE Interleukin-2.
XX Interleukin-2; lysine-depleted variant; site-directed
KW mutagenesis; human.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 43..501
FT /*tag= a
XX
XX W08905824-A.
XX
XX
XX
XX PD 29-JUN-1989.
XX
XX PF 22-DEC-1988; 88WO-US04633.
XX
XX PE 23-DEC-1987; 87US-0137043.
XX
XX (GENE) GENETICS INST INC.
XX
XX Shaw G;
XX
XX WPI: 1989-206594/28.
XX
XX DR P-PSDB: AAP90467.
XX
XX
XX New lysine depleted variants of polypeptide
PT - opt. modified with hydrophilic residues,
PT biologically active but with altered solubility, stability etc.
XX
XX
XX Disclosure; fig 1: 35pp; English.
XX
XX DNA of interleukin-2 (see corresp. AAP90467). Used in the patent
CC to create lysine depleted variants by site-directed mutagenesis, or
CC synthesis.
XX
XX Sequence 788 BP; 272 A; 146 C; 113 G; 257 T; 0 other;
SQ

Query Match 100.0%; Score 29; DB 10; Length 788;
Best Local Similarity 58.6%; Pred. No. 0.00053;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuuaagccuagggcu 29
Db :|||:|||||:|||||:|||||:
645 tatgattctttgttaagccctagggcct 673

RESULT 14
AAN40441
ID AAN40441 standard; cDNA; 794 BP.
XX
XX AAN40441;
AC
XX
XX 21-JUL-1992 (first entry)
DT
XX
XX Gene encoding polypeptide having IL-2 activity.
DE
XX
XX Interleukin.
KW
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 47..508
FT /*tag= a
FT PolYA_signal 770..775
FT /*tag= b
FT PolYA_site 788..794
FT /*tag= c
XX
XX JP59139349-A.

XX
PD 10-AUG-1984.
XX
XX 29-DEC-1982; 82JP-0230372.
PE
XX
XX 29-DEC-1982; 82JP-0230372.
PR
XX
XX (AJIN) AJINOMOTO KK.
PA (GANK-) GAN KENKYUKAI ZH.
XX
XX WPI: 1984-233967/38.
DR
XX
XX Polypeptide having interleukin-2 activity - with amino acid sequence
PT having C-terminal threonine.
PT
XX
XX Disclosure; Page 7; 16pp; Japanese.
PS
XX
XX The DNA was obt'd. from mRNA of human Leukaemia cells.
CC
XX
XX Sequence 794 BP; 276 A; 147 C; 113 G; 258 T; 0 other;
SQ

Query Match 100.0%; Score 29; DB 5; Length 794;
Best Local Similarity 58.6%; Pred. No. 0.00053;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

OY 1 uaugauucuuuuuuaagccuagggcu 29
Db :|||:|||||:|||||:|||||:
649 tatgattctttgttaagccctagggcct 677

RESULT 15
AAN90381
ID AAN90381 standard; cDNA; 794 BP.
XX
XX AAN90381;
AC
XX
XX 03-NOV-1989 (first entry)
DT
XX
XX Recombinant human interleukin-2 DNA.
DE
XX
XX Recombinant human interleukin-2 DNA; jurkat cells; leukaemia.
KW
XX
XX Homo sapiens (human).
OS
XX
XX Key Location/Qualifiers
FH sig_peptide 47..106
FT /*tag= a
FT mat_peptide 107..505
FT /*tag= b
XX
XX JP01165399-A.
XX
XX 29-JUN-1989.
PD
XX
XX 24-DEC-1982; 82JP-0292084.
PE
XX
XX 24-DEC-1982; 82JP-0292084.
PR
XX
XX (AJIN) AJINOMOTO KK.
PA (GANK-) ZH GAN KENKYUKAI.
XX
XX WPI: 1989-230772/32.
DR P-PSDB: AAP90614.
XX
XX Recombinant human interleukin-2
PT - produced by irradiating hIL-2-producing
PT jurkat cells, extracting mRNA, forming cDNA, inserting into
PT pBR322 and transforming E. coli.
XX
XX Claim 1; page 737; 16pp; Japanese.
XX
XX Recombinant human interleukin-2 (hIL-2) DNA (see AAP90614). Isolated

SQ Sequence 794 BP; 276 A; 146 C; 114 G; 258 T; 0 other;

Matches	17;	conservative	12;	Mismatches
---------	-----	--------------	-----	------------

Db 649 tatgatctctttgttaagccctagggct 677

od time: 3809 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:38:42 ; Search time 1643.83 Seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagaauuuuuuuaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est148:*
118: gb_est149:*
119: gb_est150:*
120: gb_est151:*
121: gb_est152:*
122: gb_est153:*
123: gb_est154:*
124: gb_est155:*
125: gb_est156:*
126: gb_est157:*
127: gb_est158:*
128: gb_est159:*
129: gb_est160:*
130: gb_est161:*
131: gb_est162:*
132: gb_est163:*
133: gb_est164:*
134: gb_est165:*
135: gb_est166:*
136: gb_est167:*
137: gb_est168:*
138: gb_est169:*
139: gb_est170:*
140: gb_est171:*
141: gb_est172:*
142: gb_est173:*
143: gb_est174:*
144: gb_est175:*
145: gb_est176:*
146: gb_est177:*
147: gb_est178:*
148: gb_est179:*
149: gb_est180:*
150: gb_est181:*
151: gb_est182:*
152: gb_est183:*
153: gb_est184:*
154: gb_est185:*
155: gb_est186:*
156: gb_est187:*
157: gb_est188:*
158: gb_est189:*
159: gb_est190:*
160: gb_est191:*
161: gb_est192:*
162: gb_est193:*
163: gb_est194:*
164: gb_est195:*
165: gb_est196:*
166: gb_est197:*
167: gb_est198:*
168: gb_est199:*
169: gb_est200:*
170: gb_est201:*
171: gb_est202:*
172: gb_est203:*
173: gb_est204:*
174: gb_est205:*
175: gb_est206:*
176: gb_est207:*
177: gb_est208:*
178: gb_est209:*
179: gb_est210:*
180: gb_est211:*
181: gb_est212:*
182: gb_est213:*
183: gb_est214:*
184: gb_est215:*
185: gb_est216:*
186: gb_est217:*
187: gb_est218:*
188: gb_est219:*
189: gb_est220:*

190: gb_est221:*
191: gb_est222:*
192: gb_est223:*
193: gb_est224:*
194: gb_est225:*
195: gb_est226:*
196: gb_est227:*
197: gb_est228:*
198: gb_est229:*
199: gb_est230:*
200: gb_est231:*
201: gb_est232:*
202: gb_est233:*
203: gb_est234:*
204: gb_est235:*
205: gb_est236:*
206: gb_est237:*
207: gb_est238:*
208: gb_est239:*
209: gb_est240:*
210: gb_est241:*
211: gb_est242:*
212: gb_est243:*
213: gb_est244:*
214: gb_est245:*
215: gb_est246:*
216: gb_est247:*
217: gb_est248:*
218: gb_est249:*
219: gb_est250:*
220: gb_est251:*
221: gb_est252:*
222: gb_est253:*
223: gb_est254:*
224: gb_est255:*
225: gb_est256:*
226: gb_est257:*
227: gb_est258:*
228: gb_est259:*
229: gb_est260:*
230: gb_est261:*
231: gb_est262:*
232: gb_est263:*
233: gb_est264:*
234: gb_est265:*
235: gb_est266:*
236: gb_est267:*
237: gb_est268:*
238: gb_est269:*
239: gb_est270:*
240: gb_est271:*
241: gb_est272:*
242: gb_est273:*
243: gb_est274:*
244: gb_est275:*
245: gb_est276:*
246: gb_est277:*
247: gb_est278:*
248: gb_est279:*
249: gb_est280:*
250: gb_est281:*
251: gb_est282:*
252: gb_est283:*
253: gb_est284:*
254: gb_est285:*
255: gb_est286:*
256: gb_est287:*
257: gb_est288:*
258: gb_est289:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

QY	1	aagaauuucuuuuuagaccccaagg	27	: : : :	93.1%	Score 27:	DB 4:	Length 325:
Db	167	AAAGATCTTTTGTAGCCCAAGG	193		70.4%	Pred. No. 0.13:		
		Matches 19:	Conservative 8:	Mismatches 0:	Indels 0:	Gaps 0:		

RESULT	2
LOCUS	A0267968/c
DEFINITION	A0267968 537 bp DNA GSS 27-APR-1989 RC111-73E11.TJ RcP1-11 Homo sapiens genomic clone RcP1-11-73E11, DNA sequence.
ACCESSION	A0267968
VERSION	A0267968
KEYWORDS	A0267968.1 GI:3795572 GSS.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 527) Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linhar,K., Golden,K., Berry,K., Granger,D., Suh,E., Wille,C., de Jong,P. and Venter,J.C. Use of human BAC End Sequences for Sequence-Ready Map Building Unpublished (1998)
TITLE	Other GSSs: RC111-73E11.TK
JOURNAL	Contact: Mark Adams
COMMENT	Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
Fax: 301 838 0208

Email: mdadams@tigr.org
Library available from the human BAC library RPCI-11. For BAC clones are derived from the human BAC library Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..527
/organism="Homo sapiens"
/db_xref="GDB:7527754"
/db_xref="taxon:9606"
/clone="RPCI-11-73E11"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC library"
BASE COUNT 203 a 78 c 77 g 167 t 2 others
ORIGIN

Query Match 80.0%; Score 23.2; DB 226; Length 527;
Best Local Similarity 57.1%; Pred. No. 6.5;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 2 aaagauuuuuuuaagccccaaggcu 29

Db 357 ATGATTCCTTTTGTAGCCCTAGGGGCT 330

RESULT 3

LOCUS

AQ267210 596 bp DNA GSS 27-APR-1999
DEFINITION RPCI11-73C7.7J RPCI-11 Homo sapiens genomic clone RPCI-11-73C7, DNA sequence.

ACCESSION

AQ267210
AQ267210.1 GI:3794814

VERSION

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 596)

Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Middle, C., de Jong, P., and Venter, J.C.

Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)

Other GSSs: RPCI11-73C7.TK

JOURNAL

COMMENT

TITLE

AUTHORS

REFERENCE

BIBLIOGRAPHY

CITATIONS

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

/clone="RPCI-11-73C7"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC library"
BASE COUNT 227 a 92 c 88 g 186 t 3 others
ORIGIN

Query Match 80.0%; Score 23.2; DB 226; Length 596;
Best Local Similarity 57.1%; Pred. No. 6.6;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 2 aaagauuuuuuuaagccccaaggcu 29

Db 357 ATGATTCCTTTTGTAGCCCTAGGGGCT 330

RESULT 4

LOCUS

BG365004 246 bp mRNA EST 08-MAR-2001
DEFINITION 112339 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION

BG365004
BG365004.1 GI:13254101

VERSION

KEYWORDS

SOURCE

ORGANISM

COW

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 246)

Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M.,

Bennett, G.A., Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Laegreid

, W.W. and Keefe, J.W.

Design and use of four pooled tissue normalized cDNA libraries for

EST discovery in cattle

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt. trimmed with phred

v0.980904.e. Vector identified by cross-match with the -minscore 18

and -minmatch 12 options.

PCR primers

FORWARD: AGGAACAGCATATGACCAT

BACKWARD: GTTTCACAGTACGACG

Plate: 103 row: K column: 22

Seq primer: ATTAGGACACATATAG.

location/Qualifiers

1..246

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"

/tissue_type="Pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from lymph node, ovary,

fat, hypothalamus, and pituitary."

BASE COUNT 62 a 49 c 68 g 67 t

ORIGIN

Query Match 72.4%; Score 21; DB 152; Length 246;
Best Local Similarity 51.7%; Pred. No. 53;
Matches 15; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

OY 1 aaagauuuuuuuaagccccaaggcu 29

Db 198 AAAGGTTTTTTTAAAGCCCAAGCCCT 170

RESULT 5
LOCUS A0809431/c 516 bp DNA GSS 10-AUG-1999
DEFINITION HS_3149_B1_A05_TTC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3149 Col=9 Row=B, DNA sequence.
ACCESSION A0809431
VERSION A0809431.1 GI:5728673
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 3149 row: B column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 516.
FEATURES
source Location/Qualifiers
1..516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3149 Col=9 Row=B"
/clone_1ib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 162 a 112 c 99 g 139 t 4 others
ORIGIN

Query Match 72.4%; Score 21; DB 233; Length 516;
Best Local Similarity 58.6%; Pred. No. 60;
Matches 17; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 1 aaagaucuuuuuaggaagcccaaggcu 29
||| | : : : ||||| ||| | : :
Db 90 AAACACTCTGTTGTAAGCCGCAAGTCT 62

RESULT 6
LOCUS A0239971 309 bp DNA GSS 30-SEP-1998
DEFINITION CIT-HSP-2386F6.TR.1 CIT-HSP Homo sapiens genomic clone 2386F6, DNA sequence.
ACCESSION A0239971
VERSION A0239971.1 GI:3672169
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 309)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2386F6.TR.1
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC clones end search page:
http://www.tigr.org/tcdb/hungen/Bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2386F6"
/clone_1ib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 105 a 56 c 58 g 90 t
ORIGIN

Query Match 71.7%; Score 20.8; DB 226; Length 309;
Best Local Similarity 58.3%; Pred. No. 68;
Matches 14; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

OY 6 uucuuuuuaggaagcccaaggcu 29
::: : : : ||||| ||| | : :
Db 80 TTCTATTGTAAGCCCATGGCT 57

RESULT 7
LOCUS AZ407430 646 bp DNA GSS 03-OCT-2000
DEFINITION IM0178F16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0178F16 F, DNA sequence.
ACCESSION AZ407430
VERSION AZ407430.1 GI:10531539
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 646)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb Plasmid Inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0178 row: F column: 16
Seq primer: CGTGTAAACGACGCGCAGT
Class: Plasmid inserts
High quality sequence stop: 646.
FEATURES
Location/Qualifiers

SOURCE

```

1. 646
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UUCG1M0178F16"
/clone_lib="Mouse 10Kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and 14 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      151 a      96 c      115 g      284 t
ORIGIN

```

```

Query Match      69.0%; Score 20; DB 243; Length 646;
Best Local Similarity 53.6%; Pred. No. 1.7e+02;
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```

```

Oy 2 aagauuuuuuuuagagcccaaggcu 29
||||:||||:||||:||||:||||:
Db 134 AAGATCTTTTGGAGCAGCTGAGGCT 161

```

```

RESULT 8
LOCUS AO592351/c 470 bp DNA GSS 08-JUN-1999
DEFINITION HS_5446_A1-B11-SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate-1022 Col-21 Row-C, DNA sequence.
ACCESSION AO592351
VERSION AO592351.1 GI:5024003
KEYWORDS GSS.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 470)
Mahafras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

```

```

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahafas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

```

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@edj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

```

Plate: 1022 row: C column: 21
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 470.
FEATURES
Location/Qualifiers
1..470

```

SOURCE

```

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="Plate-1022 Col-21 Row-C"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="Male"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
BASE COUNT      144 a      89 c      89 g      147 t      1 others
ORIGIN

```

```

Query Match      68.3%; Score 19.8; DB 230; Length 470;
Best Local Similarity 60.9%; Pred. No. 2e+02;
Matches 14; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy 4 gaauuuuuuuuagagcccaagg 26
||||:||||:||||:||||:||||:
Db 76 GTTCTTTTGAAGCCCAAGG 54

```

```

RESULT 9
LOCUS AO487601 679 bp DNA GSS 24-APR-1999
DEFINITION RPCI-11-242H19.JY RPCI-11 Homo sapiens genomic clone RPCI-11-242H19
, DNA sequence.
ACCESSION AO487601
VERSION AO487601.1 GI:4673475
KEYWORDS GSS.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 679)
Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter
,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Other GSSs: RPCI-11-242H19.TV
Contact: Shaying Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@igrr.org

```

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@edj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html. Seq primer: SP6

```

FEATURES
Class: BAC ends.
Location/Qualifiers
1..679

```

```

SOURCE
/organism="Homo sapiens"
/db.xref="GDB:7592730"
/db.xref="taxon:9606"
/clone="RPCI-11-242H19"
/clone_lib="RPCI-11"
/sex="Male"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;"

```

BASE COUNT 200 a 146 c 142 g 191 t
 ORIGIN

Query Match 68.3%; Score 19.8; DB 229; Length 679;
 Best Local Similarity 60.9%; Pred. No. 2.1e+02;
 Matches 14; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 4 gaucuuuuuuuagccccaag 26
 Db 501 GTTCTTTTGAAGCCCAAG 523

RESULT 10
 LOCUS AQ487546 694 bp DNA GSS 24-APR-1999
 DEFINITION RPCI-11-242F17.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-242F17
 COMMENT , DNA sequence.
 ACCESSION AQ487546
 VERSION AQ487546.1 GI:4673420
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
 J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other GSSs: RPCI-11-242F17.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
 Research Genet cs (inforesgen.com). BAC end search page:
 http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers

1..694
 /organism="Homo sapiens"
 /db_xref="GDB:7592860"
 /db_xref="taxon:9606"
 /clone="RPCI-11-242F17"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

BASE COUNT 202 a 151 c 143 g 198 t
 ORIGIN

Query Match 68.3%; Score 19.8; DB 229; Length 694;
 Best Local Similarity 60.9%; Pred. No. 2.1e+02;
 Matches 14; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 4 gaucuuuuuuuagccccaag 26
 Db 512 GTTCTTTTGAAGCCCAAG 534

RESULT 11

AQ487597
 LOCUS AQ487597 723 bp DNA GSS 24-APR-1999
 DEFINITION RPCI-11-242H15.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-242H15
 COMMENT , DNA sequence.

ACCESSION AQ487597
 VERSION AQ487597.1 GI:4673471
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
 J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other GSSs: RPCI-11-242H15.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics -
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
 Research Genet cs (inforesgen.com). BAC end search page:
 http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers

1..723
 /organism="Homo sapiens"
 /db_xref="GDB:7592726"
 /db_xref="taxon:9606"
 /clone="RPCI-11-242H15"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

BASE COUNT 215 a 150 c 153 g 205 t
 ORIGIN

Query Match 68.3%; Score 19.8; DB 229; Length 723;
 Best Local Similarity 60.9%; Pred. No. 2.1e+02;
 Matches 14; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 4 gaucuuuuuuuagccccaag 26
 Db 501 GTTCTTTTGAAGCCCAAG 523

RESULT 12
 LOCUS CNS000GB/c 1101 bp DNA GSS 03-JUN-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
 BACR3B22 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL071964
 VERSION AL071964.1 GI:4952148
 KEYWORDS GSS.
 SOURCE fruit fly.

ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 Genoscope.

BASE COUNT 93 a 69 c 52 g 136 t
ORIGIN

Query Match 67.6%; Score 19.6; DB 11; Length 350;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 16; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 aaagaucuuuuuagccccaag 26
|||||:|:::| ||| |||||
Db 192 AAAGATCTTTTGGAGGCCCAAG 217

RESULT 15

AO561664 428 bp DNA GSS 29-MAY-1999
LOCUS HS_5206.A1.C10.SPEE RPCI-11 Human Male BAC library Homo sapiens
DEFINITION genomic clone Plate=782 Col=19 Row=E, DNA sequence.
AO561664
ACCESSION AO561664.1 GI:4921135
VERSION GSS.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 428)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.bufile.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufile.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
<http://www.htsc.washington.edu>
Place: 782 row: E column: 19
Seq primer: SP6
Class: BAC ends

High quality sequence stop: 428.

FEATURES
source
1.428
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=782 Col=19 Row=E"
/clone_lib="RPCI-11 Human Male BAC library"
/sex="male"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

BASE COUNT 140 a 95 c 72 g 116 t 5 others
ORIGIN

Query Match 67.6%; Score 19.6; DB 230; Length 428;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 14; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 1 aaagaucuuuuuagccccaag 26
|||||:|:::| ||| |||||
Db 61 AAAGATCTTTTCCCAAG 86

Search completed: October 2, 2001, 04:38:47
Job time: 4069 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:00:41 ; Search time 1315.38 Seconds
(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-25

Perfect score: 29

Sequence: 1 aaagaunuuuuuuuuaagccccaagggcu 29

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Actual number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_p11:*

13: gb_p12:*

14: gb_p13:*

15: gb_p14:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_hcgo_hum:*

20: em_hcgo_in:*

21: em_hcgo_rnd:*

22: em_hcgo_hum1:*

23: em_hcgo_hum2:*

24: em_hcgo_hum3:*

25: em_hcgo_hum4:*

26: em_hcgo_hum5:*

27: em_hcgo_hum6:*

28: em_hcgo_hum7:*

29: em_hcgo_hum8:*

30: em_hcgo_in:*

31: em_hcgo_in:*

32: em_hcgo_in:*

33: em_hcgo_in:*

34: em_hcgo_in:*

35: em_hcgo_in:*

36: em_hcgo_in:*

37: em_hcgo_in:*

38: em_hcgo_in:*

39: em_hcgo_in:*

40: em_hcgo_in:*

41: em_hcgo_in:*

42: em_hcgo_in:*

43: em_hcgo_in:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_p1:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_v1:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_v11:*

59: gb_v12:*

60: gb_hc1:*

61: gb_hc2:*

62: gb_hc3:*

63: gb_hc4:*

64: gb_hc5:*

65: gb_hc6:*

66: gb_hc7:*

67: gb_hc8:*

68: gb_hc9:*

69: gb_hc10:*

70: gb_hc11:*

71: gb_hc12:*

72: gb_hc13:*

73: gb_hc14:*

74: gb_hc15:*

75: gb_hc16:*

76: gb_hc17:*

77: gb_hc18:*

78: gb_hc19:*

79: gb_hc20:*

80: gb_hc21:*

81: gb_hc22:*

82: gb_hc23:*

83: gb_hc24:*

84: gb_hc25:*

85: gb_p1:*

86: gb_p2:*

87: gb_p3:*

88: gb_p4:*

89: gb_p5:*

90: gb_p6:*

91: gb_p7:*

92: gb_p8:*

93: gb_p9:*

94: gb_p10:*

95: gb_p11:*

96: gb_p12:*

97: gb_p13:*

98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	100.0	740	95	RATIL2	M22899 Rat Interle
2	29	100.0	822	94	MMIL03	X01665 Mouse Inter
3	29	100.0	825	10	E00849	E00849 Nucleic aci
4	29	100.0	825	10	I01928	I01928 Sequence 1
5	29	100.0	825	94	MUSIL2T	K02292 Mouse Inter
6	29	100.0	868	94	MMINTL2	X73040 M. musculus
7	29	100.0	870	94	MUSIL2A03	M16762 Mouse Inter
8	29	100.0	939	45	E10514	E10514 cDNA encodi


```

E00849      E00849      825 bp      RNA      PAT      29-SEP-1997
LOCUS       Nucleic acid sequence of cDNA encoding a polypeptide possessing
DEFINITION  murine IL-2 activity.
ACCESSION   E00849
VERSION     E00849.1 GI:2169110
KEYWORDS    JP 1986119197-A/1.
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 825)
AUTHORS    Furukawa, D., Yokota, T., and Arai, K.
TITLE      CDNA CLONE FOR CODING POLYPEPTIDE SHOWING RAT INTERLEUKIN-2
JOURNAL     Patent: JP 1986119197-A 1 06-JUN-1986;
COMMENT     SHERINGU BAIOTETSUKU CORP
OS          Human [Homo sapiens]
PN          JP 1986119197-A/1
PD          06-JUN-1986
PE          04-OCT-1985 JP 1985221684
PR          05-OCT-1984 US 84 658183
PI          FURUKAWA DON RIT, YOKOTA TAKASHI, ARAI KENICHI PC
PC          C12P21/02, C07K13/00, C07K15/04, C12N1/00, C12N5/00, C12N15/00, PC
PC          G01N33/53, G01N33/577, (C12P21/02, C12R1:19), (C12P21/02, PC
PC          C12R1:865), (C12P21/02, PC
PC          C12R1:91), (C12N1/00, C12R1:19), (C12N1/00, C12R1:865), (C12N5/00,
PC          C12R1:91);
CC          strandedness: Double;
CC          topology: linear;
CC          hypothetical: No;
CC          anti-sense: No;
CC          *source: tissue-type-blood;
CC          *source: cell-type-T-cell;
CC          *source: cell-line-LB2-1 cell;
CC          *source: clone=LB2-1 clone;
FH          Key          Location/Qualifiers
FH          CDS          49..558
FH          FT          /product="polypeptide possessing murine IL-2
FT          FT          activity".
FEATURES
source          1..825
                  Location/Qualifiers
BASE COUNT     252 a 179 c 155 g 239 t
ORIGIN
Query Match    100.0%; Score 29; DB 10; Length 825;
Best Local Similarity 69.0%; Pred. No. 0.004;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
Oy 1 aaagaucuuuuuuaagcccaagggcu 29
Db 688 AAAGATCTTTTGTAAAGCCCAAGGGCT 716

RESULT 4
LOCUS       101928      825 bp ss-DNA
DEFINITION  Sequence 1 from Patent US 4798789.
ACCESSION   101928
VERSION     101928.1 GI:269761
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 825)
AUTHORS    Lee, F.D., Yokota, T., and Arai, K.
TITLE      cDNA clones coding for polypeptides exhibiting murine interleukin-2
JOURNAL     Patent: US 4798789-A 1 17-JAN-1989;

```

```

DNAX Research Institute of Molecular and Cellular Biology, Inc.,
Palo Alto, CA
FEATURES
source          1..825
                  Location/Qualifiers
BASE COUNT     252 a 179 c 155 g 239 t
ORIGIN
Query Match    100.0%; Score 29; DB 10; Length 825;
Best Local Similarity 69.0%; Pred. No. 0.004;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
Oy 1 aaagaucuuuuuuaagcccaagggcu 29
Db 688 AAAGATCTTTTGTAAAGCCCAAGGGCT 716

RESULT 5
LOCUS       MUSTL2T      825 bp      mRNA      ROD      12-JUN-1993
DEFINITION  Mouse interleukin-2 mRNA, complete cds.
ACCESSION   K02292
VERSION     K02292.1 GI:198330
KEYWORDS    T-cell growth factor; interleukin; interleukin 2.
SOURCE      Mouse (T-cell line LB2-1, from T-cell line C57BL/6), cDNA to mRNA,
            clone MF-1.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 825)
AUTHORS    Yokota, T., Arai, N., Lee, F., Remnick, D., Mosmann, T., and Arai, K.-I.
TITLE      Use of a cDNA expression vector for isolation of mouse interleukin
JOURNAL     2 cDNA clones: Expression vector for isolation of mouse interleukin
MEDLINE     Transfection of monkey cells
PROC        Proc. Natl. Acad. Sci. U.S.A. 82, 68-72 (1985)
FEATURES
source          1..825
                  Location/Qualifiers
sig_peptide    49..108
                  /organism="Mus musculus"
                  /db_xref="taxon:10090"
CDS            49..558
                  /note="interleukin 2 signal peptide"
                  /note="interleukin 2 prepeptide"
                  /codon_start=1
                  /protein_id="AAA39289.1"
                  /translation="MYSMQLASCVTLLVLVNSAPTSSTSSSTAFAEQ00000000
                  OOHLEFOLMDLOELISRMENYRNKLPLPFPKFLPPOATFELKDLOCLEDELGLRH
                  VIDLQSKSFQLEDAENFTSNRTVTVKLSGSDNTECFQPDDESATVVDPLRRMTAFC
                  OSITSPQ"
mat_peptide    109..555
                  /note="interleukin 2"
BASE COUNT     252 a 179 c 155 g 239 t
ORIGIN
Query Match    100.0%; Score 29; DB 94; Length 825;
Best Local Similarity 69.0%; Pred. No. 0.004;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
Oy 1 aaagaucuuuuuuaagcccaagggcu 29
Db 688 AAAGATCTTTTGTAAAGCCCAAGGGCT 716

RESULT 6
LOCUS       MMINTLK2      868 bp      mRNA      ROD      17-FEB-1997
DEFINITION  M.musculus mRNA for interleukin-2.
ACCESSION   X73040
VERSION     X73040.1 GI:397468

```



```

RL AJINOMOTO CO INC, JAPAN FOUND CANCER RES.
XX
CC OS Mus sp. (mouse)
CC PN JP 1996000272-A/1
CC PD 09-JAN-1996 JP 1995180603
CC PI KASHIMA SHINICHI, HAMURO JUNJI, TANIGUCHI KOREAKI,
CC PI SUGANO HARUO
CC PC C12N15/09,C07H21/04,C12N1/21,C12N5/10,C12P21/02//A61K35/12,
CC PC A61K35/74
CC PC A61K38/00,C07K14/55,(C12N1/21,C12R1:19),(C12P21/02,C12R1:19),
CC PC (C12P21/02,
CC PC C12R1:91);
CC PC C12R1:91);
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC Key
CC FH Location/Qualifiers
CC FT source
CC FT 1..939
CC FT /organism="Mus sp."
CC FT /cell_line="LBRM-33"
CC FT /tissue_type="lymphoma"
CC FT 5'UTR
CC FT 1..47
CC FT sig_peptide
CC FT 48..110
CC FT /product="signal peptide of mouse
CC FT interleukin-2"
CC FT 111..554
CC FT mat_peptide
CC FT 48..557
CC FT /product="mouse interleukin-2"
CC FT CDS
CC FT 48..557
CC FT /product="mouse interleukin-2"
CC FT 3'UTR
CC FT 558..939
XX
FH Key
FH Location/Qualifiers
FH source
FH 1..939
FH /db_xref="taxon:10095"
FH /organism="Mus sp."
FT
FT
FT
FT
XX
SQ Sequence 939 BP; 287 A; 202 C; 173 G; 277 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 29; DB 45; Length 939;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagaunuuuuuuuagaccaccaagagcu 29
|||||:|||||:|||||:|||||:|||||:
687 AAAGATTCTTTTGTAAAGCCCAAGGCT 715

RESULT 9
LOCUS MM104 939 bp mRNA ROD 30-MAR-1995
DEFINITION Mouse mRNA for interleukin-2 (IL-2).
ACCESSION X01772 K02797
VERSION X01772.1 GI:52663
KEYWORDS interleukin; signal peptide; tandem repeat.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Kashima,N., Nishi-Takaoka,C., Fujita,T., Taki,S., Yamada,G.,
Hamuro,J. and Taniguchi,T.
TITLE Unique structure of murine interleukin-2 as deduced from cloned
CDNAs
JOURNAL Nature 313 (6001), 402-404 (1985)
MEDLINE 8511148
COMMENT Data kindly reviewed (19-FEB-1986) by G. Yamada.
FEATURES
source
1..939
/organism="Mus musculus"

```

```

CDS
/db_xref="taxon:10090"
48..557
/codon_start=1
/product="interleukin-2"
/protein_id="CAA25909.1"
/db_xref="GI:758159"
/db_xref="MCD:MGI:96548"
/db_xref="SWISS-PROT:P04351"
/translation="MYSMLASGVTVTVLVNLSAPTSSTSSSTAFAA000000000
VLDLQSKSFQLEDAENFTSNIRVTYVTKLGSNDTFECQFDDESATVDFLRWIAFC
OSLISTSPQ"
48..107
/product="signal peptide (aa -21 to -1)"
108..554
/product="interleukin-2 (aa 1-133)"
repeat_region
/feature="12 tandem repeats of CAG"
/note="12 tandem repeats of CAG"
BASE COUNT 287 a 202 c 173 g 277 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 29; DB 94; Length 939;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagaunuuuuuuuagaccaccaagagcu 29
|||||:|||||:|||||:|||||:|||||:
Db 687 AAAGATTCTTTTGTAAAGCCCAAGGCT 715

RESULT 10
LOCUS AF195956 24900 bp DNA ROD 04-MAY-2000
DEFINITION Mus musculus interleukin-2 (Il2) gene, complete cds; and testis
nuclear RNA-binding protein (Tenr) gene, partial cds.
ACCESSION AF195956
VERSION AF195956.1 GI:6941952
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Lyons,P.A., Arnltage,N., Argentina,F., Denny,P., Hill,N.J.,
Lord,C.J., Wilusz,M.B., Peterson,L.B., Wicker,L.S. and Todd,J.A.
TITLE Congenic mapping of the type 1 diabetes locus, Id3, to a 780-kb
region of mouse chromosome 3: identification of a candidate segment
of ancestral DNA by haplotype mapping
JOURNAL Genome Res. 10 (4), 446-453 (2000)
MEDLINE 20243845
PUBMED 10779485
REFERENCE
AUTHORS Lyons,P.A., Arnltage,N., Lord,C.J., Denny,P., Hill,N.J.,
Podolin,P.L., Peterson,L.B., Wicker,L.S. and Todd,J.A.
TITLE Direct Substitution
JOURNAL Submitted (19-OCT-1999) Medical Genetics, University of Cambridge,
Cambridge Institute for Medical Research, Addenbrooke's Hospital,
Cambridge CB2 2XY, UK
FEATURES
source
1..24900
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="3"
/clone="mC1h5"
/complement(385..509)
/note="SINE"
/rpt_family="B1"
/rpt_type="dispersed"
2786..2853
/note="SINE"
/rpt_family="MIR"
repeat_region
repeat_region
repeat_region

```

```

/db_xref="GI:6941954"
/translation="SPEKSLTSMASRLCRAMLSRENTLAKKAETDLDLEARYTHAAK
CMGSPQDEAKALKATILQQDGYSWIKSFCTIDFSH"
complement(19651..20127)
/note="LINE"
/rpt_family="I1"
/rpt_type-dispersed
complement(20958..21142)
/note="LINE"
/rpt_family="I1"
/rpt_type-dispersed
complement(22450..22619)
/note="LINE"
/rpt_family="I1"
/rpt_type-dispersed
complement(24776..24898)
/note="SINE"
/rpl_family="B2"
/rpl_type-dispersed
complement(4720 c 4896 g 7278 t 58 others
BASE COUNT 7948 a 4720 c 4896 g 7278 t 58 others
ORIGIN
Query Match 100.0%; Score 29; DB 94; Length 24900;
Best Local Similarity 69.0%; Prd. No. 0.0068;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Caps 0;
QY 1 aaagaucuuuuuguaagcccaaggcgu 29
|||||:::.....:|||||||
Db 9497 AAAGATTCTTTGTGAAGCCCAAGGCT 9525

RESULT 11
LOCUS BOVIL2 764 bp mRNA MAM 27-APR-1993
DEFINITION Bovine interleukin 2 (IL-2) mRNA, complete cds.
ACCESSION M12791
VERSION M12791.1 GI:163204
KEYWORDS interleukin 2.
SOURCE Bovine lymph node, CDNA to mRNA, clone pBIL2-4.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 764)
Carettili,D.P., McKereghan,K., Larsen,A., Cantrell,M.A.,
Anderson,D., Gallis,S., Cosman,D. and Baker,P.E.
Cloning, sequence, and expression of bovine interleukin 2
Proc. Natl. Acad. Sci. U.S.A. 83, 3223-3227 (1986)
86205869
Draft entry and clean copy sequence for [1] kindly provided by
D.Cerretti, 12-AUG-1986.
There is probably only one copy of the interleukin 2 gene in the
bovine genome.
location/Qualifiers
1..764
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="lymph"
<1..764
/gene="IL-2"
1..764
/gene="IL-2"
18..77
/gene="IL-2"
18..485
/gene="IL-2"
/note="prepeptide"
/codon_start=1

```

/product="interleukin 2"
/protein_id="AA30586.1"
/db_xref="GI:163205"
/translation="MYKIQLSIALTLAVANGAPTSSGTGNTKEVKSLLDLOL
LEKVNPEKLSRMHDFYKYNATELKHLCLEELKLEEVNLAPSKNLNPR
ETKDSMDNIRKIVLELQSGSTFCEYDQATVNAVEFLNKKWTFQOSYSTMT"
78..482
/gene="IL-2"
/product="interleukin 2"

BASE COUNT 257 a 133 c 123 g 251 t
ORIGIN 80 bp upstream of Hg1AI site.

Query Match 85.5%; Score 24.8; DB 7; Length 764;
Best Local Similarity 60.7%; Pred. No. 0.38;
Matches 17; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 2 aagaauuuuuuuaagccccaaggcu 29
|||||:|||||:|||||:|||||:
P 620 AAGATTCTTTTGTAGCCCTACGGCT 647

RESULT 12
LOCUS CF028141 773 bp mRNA MAM 29-JUN-1995
DEFINITION Canis familiaris interleukin-2 mRNA, complete cds.
ACCESSION U28141
VERSION U28141.1 GI:881935
KEYWORDS
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 773)
Somberg, R.L., Tipold, A., Henthorn, P.S. and Felsburg, P.J.
REFERENCE
Unpublished
2 (bases 1 to 773)
Somberg, R.L.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-1995) Richard L. Somberg, Clinical Studies,
University of Pennsylvania, 3850 Spruce Street, Philadelphia, PA
19104, USA

FEATURES
source Location/Qualifiers
1..773
/organism="Canis familiaris"
/db_xref="taxon:9615"
39..506
/codon_start=1
/product="interleukin-2"
/protein_id="AA68969.1"
/db_xref="GI:881936"
/translation="MYKQILSCIALTLIVANSAPITSSGTGNTKEEQOEOLLDIQL
LLNGVNNENPQLSRMLTFKFTPKKATFTHLOCLAEELKLEEVNLAPSKNHLT
DTKELISNNNVTLTKGSETYNCEYDEDTATTFEFLNKKWTFQOSIFSTLT"
773
/note="38 A nucleotides"

BASE COUNT 265 a 145 c 118 g 245 t
ORIGIN

Query Match 85.5%; Score 24.8; DB 7; Length 773;
Best Local Similarity 60.7%; Pred. No. 0.38;
Matches 17; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 2 aagaauuuuuuuaagccccaaggcu 29
|||||:|||||:|||||:|||||:
Db 641 AAGATTCTTTTGTAGCCCTACGGCT 668

RESULT 13
LOCUS I82323 209 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 8 from patent US 5712126.

ACCESSION I82323
VERSION I82323.1 GI:3210620
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 209)
Weissman, S.M. and Prashar, Y.
AUTHORS
TITLE Analysis of gene expression by display of 3'-end restriction
fragments of cDNA
JOURNAL Patent: US 5712126-A 8 27-JAN-1998;
FEATURES
source Location/Qualifiers
1..209
/organism="unknown"

BASE COUNT 56 a 30 c 33 g 90 t
ORIGIN

Query Match 80.0%; Score 23.2; DB 10; Length 209;
Best Local Similarity 57.1%; Pred. No. 1.7;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 aagaauuuuuuuaagccccaaggcu 29
|||||:|||||:|||||:|||||:
Db 34 AAGATTCTTTTGTAGCCCTACGGCT 61

RESULT 14
LOCUS G06364 292 bp DNA STS 19-OCT-1995
DEFINITION human STS WI-7035.
ACCESSION G06364
VERSION G06364.1 GI:859609
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene
collection.
ORGANISM Homo sapiens
Eukaryota; Eukaryota; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Vertebrata; Mammalia; Theria; Eutheria; Archonta; Primates;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivora; Homnidae; Homo.
1 (bases 1 to 292)
Hudson, T.
REFERENCE
Whitehead Institute/MIT Center for Genome Research; Physically
Mapped ESTs
JOURNAL Unpublished (1995)
COMMENT

Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: TAATTAGTCTCCACTTAAC
Primer B: ATTTGGGATTAATAGTGAACCA
STS size: 200
PCR Profile:

Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:

Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:

MgCl₂: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 9.3

FEATURES Prepared with primer pairs derived from V00564 -- Unigene.

Source Location/Qualifiers
STS 1..292 /organism="Homo sapiens"
primer_bind 1..200
primer_bind 1..25
BASE COUNT 92 a 32 c 32 g 124 t 12 others
ORIGIN

Query Match 80.0%; Score 23.2; DB 54; Length 292;
Best Local Similarity 57.1%; Pred. No. 1.8;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

2 aagaucuuuuuuuagaccaccccaaggcu 29
||||:||||:||||:||||:||||:
142 ATGATCTCTTTGTGAAGCCCTAGGGGCT 169

RESULT 15
A17717 A17717 689 bp mRNA PAT 17-MAY-1994
LOCUS
DEFINITION Plasmid PTG 26 mRNA for human IL-2.
ACCESSION A17717
VERSION A17717.1 GI:513949
KEYWORDS
SOURCE unidentified.
ORGANISM Plasmid unidentified
unclassified.

REFERENCE 1 (bases 1 to 689)
AUTHORS Kieny M.P., Sandermeier P. and Lecocq J.P.
TITLE Expression of human IL-2 in mammalian cells by a recombinant pox virus
JOURNAL Patent: EP 0206939-A 1 30-DEC-1986;
TRANSGENE S.A

FEATURES Location/Qualifiers
Source 1..689
/organism="unidentified"
/plasmid="PTG26"
/db_xref="taxon:32644"
1..384
/partial
/codon_start=1
/product="human interleukin 2"
/protein_id="CAA01347.1"
/db_xref="GI:513950"
/translation="TKKTOLEHLDLQMLGNNKPKLTRMLTFKPYMKPKKA
TELKHLQCLEELKPLEEVLNIAOSKNFHLRPRLLISINIVIVLELGSSETFMCEYA
DETATVEFLNRWTFQCSITSTLT"

BASE COUNT 251 a 117 c 99 g 222 t
ORIGIN

Query Match 80.0%; Score 23.2; DB 9; Length 689;
Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 aagaucuuuuuuuagaccaccccaaggcu 29
||||:||||:||||:||||:||||:
Db 526 ATGATCTCTTTGTGAAGCCCTAGGGGCT 553

Search completed: October 2, 2001, 05:00:42
Job time: 5329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:03:57 ; Search time 122.48 Seconds
(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagaauuuuuuuaagcccaaggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
tal number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_0601:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	29	AAA70829	Molecular interact
2	29	100.0	29	AAA70830	Molecular interact
3	29	100.0	42	AAA71115	Molecular interact
4	29	100.0	42	AAA71116	Molecular interact
5	29	100.0	42	AAA71120	Molecular interact
6	29	100.0	42	AAA71121	Molecular interact
7	29	100.0	42	AAA71128	Molecular interact
8	29	100.0	42	AAA71129	Molecular interact
9	29	100.0	825	AA60149	Sequence encoding
10	29	100.0	940	AA60906	Sequence encodes m
11	28	96.6	45	AAA70825	Molecular interact

12	28	96.6	45	AAA70826	Molecular interact
13	28	96.6	46	AAA71088	Molecular interact
14	28	96.6	46	AAA71089	Molecular interact
15	28	96.6	46	AAA71090	Molecular interact
16	28	96.6	46	AAA71105	Molecular interact
17	28	96.6	46	AAA71106	Molecular interact
18	28	96.6	46	AAA71107	Molecular interact
19	24.8	85.5	42	AAA71113	Molecular interact
20	24.8	85.5	42	AAA71118	Molecular interact
21	24.8	85.5	42	AAA71126	Molecular interact
22	23.8	82.1	46	AAA71085	Molecular interact
23	23.8	82.1	46	AAA71103	Molecular interact
24	23.2	80.0	29	AAA70828	Molecular interact
25	23.2	80.0	42	AAA71123	Molecular interact
26	23.2	80.0	42	AAA71131	Molecular interact
27	23.2	80.0	209	AA61453	IL-2 CDNA 3'end.
28	23.2	80.0	209	AA6259456	Interleukin-2 frag
29	23.2	80.0	698	AA550132	DNA sequence encod
30	23.2	80.0	698	AA60101	Sequence of human
31	23.2	80.0	722	AA60144	Sequence encoding
32	23.2	80.0	769	AA601764	Human Interleukin-
33	23.2	80.0	784	AA60046	Sequence encoding
34	23.2	80.0	784	AA60787	Cloned sequence en
35	23.2	80.0	784	AA60102	Sequence of human
36	23.2	80.0	788	AA60253	Interleukin-2. Ho
37	23.2	80.0	794	AA60441	Gene encoding poly
38	23.2	80.0	794	AA60381	Recombinant human
39	23.2	80.0	800	AA60219	Sequence encoding
40	23.2	80.0	801	AA60031	Sequence of Interl
41	23.2	80.0	801	AA60057	Cloned human Inter
42	23.2	80.0	801	AA60254	Sequence encoding
43	23.2	80.0	802	AA60279	DNA sequence conta
44	23.2	80.0	810	AA60840	Sequence encoding
45	23.2	80.0	844	AA620965	Human low adenosin

ALIGNMENTS

RESULT	1
AAA70829	standard; RNA: 29 BP.
ID	AAA70829
XX	
AC	AAA70829;
XX	
DT	27-APR-2001 (first entry)
XX	
DE	Molecular interaction site RNA #29.
XX	
KW	Modulator; identification; molecular interaction; virtual library; ss.
XX	
OS	Mus sp.
XX	
PN	W0958947-A2.
XX	
PD	18-NOV-1999.
XX	
PF	12-MAY-1999; 99WO-US10361.
XX	
PR	12-MAY-1998; 98US-0076404.
XX	
PR	12-MAY-1998; 98US-0085092.
XX	
PA	(ISIS-) ISIS PHARM INC.
PI	Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX	Hofstadler S, McNeill J;
XX	WPI: 2000-086439/07.
PT	Identifying compounds which modulate activity of target biomolecules,
XX	used to provide compounds which can be used as pharmacological,
XX	agricultural and industrial compounds -
XX	

PS Claim 235; Page 235; 405bp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCUUAACAGAAAAC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaucuuuuuuuagagcccaagggcu 29
Db 1 aaagaucuuuuuuuagagcccaagggcu 29
|||||

RESULT 2
AAAT0830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
PM Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
XX
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX
PI Hofstadler S, McNeil J;
XX
DR WPI: 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
XX
PT used to provide compounds which can be used as pharmacological,
XX
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405bp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCUUAACAGAAAAC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaucuuuuuuuagagcccaagggcu 29
Db 1 aaagaucuuuuuuuagagcccaagggcu 29
|||||

RESULT 3
AAAT1115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KM Modulator; identification; molecular interaction; virtual library; ss.
XX
OS unidentified.
XX
PN WO958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
XX
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX
PI Hofstadler S, McNeil J;
XX
DR WPI: 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
XX
PT used to provide compounds which can be used as pharmacological,
XX
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405bp; English.

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUGUUAUACGAAAUUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;

Best Local Similarity 100.0%; Pred. No. 0.00038;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagaauuuuuuuuaagccccaagggu 29

|||||

Db 4 aaagaauuuuuuuuaagccccaagggu 32

RESULT 4

AAAT71116 standard; RNA; 42 BP.

XX AAA71116;

DT 27-APR-2001 (first entry)

DE Molecular interaction site RNA #192.

XX Modulator; identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

PD 18-NOV-1999.

PF 12-MAY-1999; 99WO-US10361.

PR 12-MAY-1998; 98US-0076404.

PR 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hofstadler S, McNeil J;

DR WPI; 2000-086439/07.

PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -

PS Example 7; Figure 122; 405bp; English.

CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUGUUAUACGAAAUUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;

Best Local Similarity 100.0%; Pred. No. 0.00038;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagaauuuuuuuuaagccccaagggu 29

|||||

Db 4 aaagaauuuuuuuuaagccccaagggu 32

RESULT 5

AAAT71120 standard; DNA; 42 BP.

XX AAA71120;

DT 27-APR-2001 (first entry)

DE Molecular interaction site DNA #126.

XX Modulator; identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

PD 18-NOV-1999.

PF 12-MAY-1999; 99WO-US10361.

PR 12-MAY-1998; 98US-0076404.

PR 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hofstadler S, McNeil J;

DR WPI; 2000-086439/07.

PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCAGUUUACGAAAAAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SO Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaauuuuuuuuuaagcccaaggcu 29
DB 4 aaagaauuuuuuuuuaagcccaaggcu 32
|||||

RESULT 8
AAAT71129
ID AAAT71129 standard; RNA; 42 BP.
XX
AC AAAT71129;
XX
DT 27-APR-2001 (first entry)
XX
NE Molecular interaction site RNA #198.
XX Modulator; identification; molecular interaction; virtual library; ss.
XX Unidentified.
XX
OS
XX
PN WO958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
XX
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
XX
PI Hotstadler S, McNeill J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACACAUAAUUCAGUUUACGAAAAAUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SO Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaauuuuuuuuuaagcccaaggcu 29
DB 4 aaagaauuuuuuuuuaagcccaaggcu 32
|||||

RESULT 9
AAN60149
ID AAN60149 standard; cDNA; 825 BP.
XX
AC AAN60149;
XX
DT 13-SEP-1991 (first entry)
XX
DE Sequence encoding polypeptide murine interleukin-2 (IL-2) activity.
XX
KM T-cell biology; immune response; ss.
XX
OS
XX
PN Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 49..558
FT /*tag= a
XX
XX
FN BP177357-A.
XX
PD 09-APR-1986.
XX
PF 03-OCT-1985; 85EP-0307094.
XX
PR 05-OCT-1984; 84US-0658183.
XX
PR 17-DEC-1986; 86US-0942600.
XX
PA (SCHE) SCHERING BIOTECH CO.
XX (DNAX-) DNAX RES INST MOLEC.
XX
PI Lee FD, Yokota T, Arai K;
XX
PI
XX
DR WPI; 1986-095692/15.
DR P-PSDB; AAP60218.
XX

DB 18 aaagaauuuuuuuaagccccaaggc 45

RESULT 12

AAA70826
ID AAA70826 standard; RNA; 45 BP.

XX AAA70826;

DT 27-APR-2001 (first entry)

DE Molecular interaction site RNA #26.

KW Modulator; identification; molecular interaction; virtual library; ss.

OS Rattus sp.

PN W09958947-A2.

18-NOV-1999.

12-MAY-1999; 99WO-US10361.

12-MAY-1998; 98US-0076404.

12-MAY-1998; 98US-0085092.

(ISIS-) ISIS PHARM INC.

Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeill J;

WPI: 2000-086439/07.

Identifying compounds which modulate activity of target biomolecules,
used to provide compounds which can be used as pharmacological,
agricultural and industrial compounds -

Claim 222; Page 232; 405pp; English.

This invention describes a novel method for identifying compounds which
modulate the activity of a target biomolecule. The method uses
3-dimensional representations of the biomolecule and a library of
compounds and comprises (a) identifying at least one molecular
interaction site of the target RNA; (b) generating in silico a virtual
library of compounds predicted or calculated to interact with the
molecular interaction site; and (c) comparing 3-dimensional (3-D)
representations of the target RNA with members of the virtual library of
compounds to generate a hierarchy of the compounds ranked in accordance
with their respective ability to form physical interactions with the
molecular interaction site. The method also describes (1) RNA comprising
a joined sequence of at least 24 nucleotides but not more than 70
nucleotides and having secondary structure defined by: (a) 3 nucleotides
forming a first side of a first double stranded (ds) region; (b) 2
nucleotides forming a first side of an internal loop region; (c) 4
nucleotides forming a first side of a second ds region; (d) 4 or 5
nucleotides forming an end loop region; (e) 4 nucleotides forming a
second side of the second ds region; (f) 4 nucleotides forming a
second side of the internal loop region; and (g) 3 nucleotides forming a second
side of the first ds region; (2) a purified and isolated RNA fragment
comprising the human sequence UUUACACAUUAUCUGUUUCAGAAAUAUC (11). The
methods and products can be used for identifying agents which modulate
the activity of biomolecules, particularly RNA. Such agents can be used
as pharmaceutical, agricultural or industrial compounds.

Sequence 45 BP; 14 A; 7 C; 9 G; 15 U; 0 other;

Query Match 96.6%; Score 28; DB 21; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 aaagaauuuuuuuaagccccaaggc 28
|||||

DB 18 aaagaauuuuuuuaagccccaaggc 45

RESULT 13

AAA71088
ID AAA71088 standard; DNA; 46 BP.

XX AAA71088;

DT 27-APR-2001 (first entry)

DE Molecular interaction site DNA #111.

KW Modulator; identification; molecular interaction; virtual library; ss.

OS unidentified.

PN W09958947-A2.

18-NOV-1999.

12-MAY-1999; 99WO-US10361.

12-MAY-1998; 98US-0076404.

12-MAY-1998; 98US-0085092.

(ISIS-) ISIS PHARM INC.

Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeill J;

WPI: 2000-086439/07.

Identifying compounds which modulate activity of target biomolecules,
used to provide compounds which can be used as pharmacological,
agricultural and industrial compounds -

Example 7; Figure 121; 405pp; English.

This invention describes a novel method for identifying compounds which
modulate the activity of a target biomolecule. The method uses
3-dimensional representations of the biomolecule and a library of
compounds and comprises (a) identifying at least one molecular
interaction site of the target RNA; (b) generating in silico a virtual
library of compounds predicted or calculated to interact with the
molecular interaction site; and (c) comparing 3-dimensional (3-D)
representations of the target RNA with members of the virtual library of
compounds to generate a hierarchy of the compounds ranked in accordance
with their respective ability to form physical interactions with the
molecular interaction site. The method also describes (1) RNA comprising
a joined sequence of at least 24 nucleotides but not more than 70
nucleotides and having secondary structure defined by: (a) 3 nucleotides
forming a first side of a first double stranded (ds) region; (b) 2
nucleotides forming a first side of an internal loop region; (c) 4
nucleotides forming a first side of a second ds region; (d) 4 or 5
nucleotides forming an end loop region; (e) 4 nucleotides forming a
second side of the second ds region; (f) 4 nucleotides forming a
second side of the internal loop region; and (g) 3 nucleotides forming a second
side of the first ds region; (2) a purified and isolated RNA fragment
comprising the human sequence UUUACACAUUAUCUGUUUCAGAAAUAUC (11). The
methods and products can be used for identifying agents which modulate
the activity of biomolecules, particularly RNA. Such agents can be used
as pharmaceutical, agricultural or industrial compounds.

Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;

Query Match 96.6%; Score 28; DB 21; Length 46;
Best Local Similarity 71.4%; Pred. No. 0.0011;
Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

1 aaagaauuuuuuuaagccccaaggc 28
|||||

Db 19 aaagattcttttgaagcccaaggc 46

RESULT 14

AAA71089
ID AAA71089 standard; DNA: 46 BP.

XX
AC AAA71089;

XX 27-APR-2001 (first entry)

XX Molecular interaction site DNA #112.

DE Modulator: identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hofstadler S, McNeil J;

XX WPI: 2000-086439/07.

PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -

XX Example 7; Figure 121; 405pp; English.

XX This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUACUGUUGUACGAAAAUUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;

Query Match 96.6%; Score 28; DB 21; Length 46;

Best Local Similarity 71.4%; Pred. No. 0.0011;

Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaauuuuuuuaagcccaaggc 28
|||||:|||||:|||||:|||||

Db 19 aaagattcttttgaagcccaaggc 46

RESULT 15

AAA71090
ID AAA71090 standard; DNA: 46 BP.

XX
AC AAA71090;

XX 27-APR-2001 (first entry)

XX Molecular interaction site DNA #113.

DE Modulator: identification; molecular interaction; virtual library; ss.

XX Unidentified.

XX WO958947-A2.

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10361.

XX 12-MAY-1998; 98US-0076404.

XX 12-MAY-1998; 98US-0085092.

XX (ISIS-) ISIS PHARM INC.

XX Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;

PI Hofstadler S, McNeil J;

XX WPI: 2000-086439/07.

PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -

XX Example 7; Figure 121; 405pp; English.

XX This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUACUGUUGUACGAAAAUUC (11). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.

XX Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;

Query Match 96.6%; Score 28; DB 21; Length 46;

Best Local Similarity 71.4%; Pred. No. 0.0011;

Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagaauuuuuuuaagcccaaggc 28
|||||:|||||:|||||:|||||

Db 19 aaagattctttgttaagccccaagggc 46

Search completed: October 2, 2001, 05:03:57
Job time: 3809 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:01:49 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-25

Sequence: 1 aagaauuuuuuuuuaagccccaaggagcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Cal number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.2	80.0	209	1	US-08-510-032A-8
2	23.2	80.0	209	3	US-08-688-514-8
3	23.2	80.0	801	6	5314995-8
4	23.2	80.0	8491	2	US-08-757-439-1
5	17.6	60.7	72928	3	US-09-009-913-1
6	17.4	60.0	51259	3	US-08-781-891-209
7	17.4	60.0	51259	3	US-08-781-891-209
8	17.2	59.3	2287	1	US-08-222-619-1
9	17.2	59.3	2287	4	US-08-221-767-23
10	17.2	59.3	2287	5	PCR-US95-04075-1
11	16.8	57.9	610	2	US-08-928-926A-2
12	16.8	57.9	610	3	US-09-212-149-2
13	16.6	57.2	711	4	US-08-998-416-900
14	16.6	57.2	2446	3	US-08-834-306-17
15	16.6	57.2	2446	4	US-08-993-674A-17
16	16.4	56.6	2311	2	US-08-712-709-6
17	16.4	56.6	2311	3	US-09-111-444-6
18	16.4	56.6	2311	4	US-09-541-228-6
19	16.2	55.9	1201	1	US-09-286-805-1
20	16.2	55.9	3383	1	US-07-707-367-1
21	16.2	55.9	3846	4	US-08-845-161A-5
22	16.2	55.9	3846	4	US-09-270-751-5
23	16.2	55.9	24417	2	US-08-846-762-1
24	16.2	55.9	72928	3	US-09-009-913-1
25	16.2	55.2	1803	3	US-08-458-922-2
26	16.2	55.2	6911	1	US-08-311-174-4
27	15.8	54.5	325	2	US-08-483-695-38

28	15.8	54.5	325	2	US-07-965-285-38	Sequence 38, Appl
29	15.8	54.5	325	4	US-08-487-231-38	Sequence 38, Appl
30	15.8	54.5	325	4	US-09-201-912-38	Sequence 38, Appl
31	15.8	54.5	743	4	US-08-821-994-38	Sequence 86, Appl
32	15.8	54.5	1102	1	US-08-821-994-86	Sequence 1, Appl
33	15.8	54.5	1207	1	US-08-460-806-1	Sequence 1, Appl
34	15.8	54.5	1307	4	US-08-325-630-1	Sequence 1, Appl
35	15.8	54.5	1333	4	US-09-142-551A-1	Sequence 61, Appl
36	15.8	54.5	1390	4	US-08-821-994-61	Sequence 63, Appl
37	15.8	54.5	1434	4	US-08-821-994-62	Sequence 63, Appl
38	15.8	54.5	1441	4	US-08-821-994-63	Sequence 7, Appl
39	15.8	54.5	2646	1	US-08-365-189-7	Sequence 3, Appl
40	15.8	54.5	2688	1	US-08-088-633-3	Sequence 3, Appl
41	15.8	54.5	2688	1	US-08-245-756-3	Sequence 3, Appl
42	15.8	54.5	2688	1	US-08-441-750-3	Sequence 3, Appl
43	15.8	54.5	2688	2	US-08-441-751-3	Sequence 3, Appl
44	15.8	54.5	2688	5	PCR-US92-02521-3	Sequence 3, Appl
45	15.8	54.5	2696	1	US-07-961-522-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-510-032A-8
; Sequence 8, Application US/08510032A
; Patent No. 5712126
; GENERAL INFORMATION:
; APPLICANT: Sherman Weissman and Yarlindra Prashar
; TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,032A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-510-032A-8

Query Match 80.0%; Score 23.2; DB 1; Length 209;
Best Local Similarity 57.1%; Pred. No. 0.051;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 2 aagaauuuuuuuaagccccaaggagcu 29
DB 34 ATGATTCTTTTGTAGCCCTAGGCGCT 61

RESULT 2
US-08-688-514-8
Sequence 8, Application US/06688514
Patent No. 6010850
GENERAL INFORMATION:
APPLICANT: Sherman Weissman and Yashindra Prashar
TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
NUMBER OF SEQUENCES: end Restriction Fragments of cDNA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,514
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: Yale
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
APPLICATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-688-514-8

Query Match 80.0% Score 23.2; DB 3; Length 209;
Best Local Similarity 57.1% Pred. No. 0.051;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 aagaucuuuuuuaagcccaaggcu 29
| |||:|||||:||||| |||||
34 ATGATCTTTTGTAGCCCTAGGGCT 61

RESULT 3
5314995-8
Patent No. 5314995
APPLICANT: FELL, HENRY P.; GAYLE, MARGIT A.
TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
BASED FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/468,390
FILING DATE: 22-JAN-1990
SEQ ID NO: 8:
LENGTH: 801
5314995-8

Query Match 80.0% Score 23.2; DB 6; Length 801;
Best Local Similarity 57.1% Pred. No. 0.063;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 aagaucuuuuuuaagcccaaggcu 29
| |||:|||||:||||| |||||

Db 651 atgattctttgtagccctagggct 678

RESULT 4
US-08-757-439-1/c
Sequence 1, Application US/08757439
Patent No. 5866371
GENERAL INFORMATION:
APPLICANT: BADZIONG, Werner
APPLICANT: HABERMANN, Paul
APPLICANT: MOELLER, Joerg
APPLICANT: ARETZ, Werner
TITLE OF INVENTION: PROCESS FOR USING THE YEAST ADH II
PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
PROTEINS IN HIGH YIELDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,439
FILING DATE: 27-NOV-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19544233.4
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 18748/303/HOCE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8491 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-757-439-1

Query Match 80.0% Score 23.2; DB 2; Length 8491;
Best Local Similarity 57.1% Pred. No. 0.094;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 aagaucuuuuuuaagcccaaggcu 29
| |||:|||||:||||| |||||
Db 6653 ATGATCTTTTGTAGCCCTAGGGCT 6626

RESULT 5
US-09-009-913-1/c
Sequence 1, Application US/09009913
Patent No. 6087485
GENERAL INFORMATION:
APPLICANT: Axxs Pharmaceuticals, Inc.
TITLE OF INVENTION: Asthma Related Genes
NUMBER OF SEQUENCES: 339
CORRESPONDENCE ADDRESS:
ADDRESS: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

Query Match 60.7%; Score 17.6; DB 3; Length 72928;
Best Local Similarity 45.8%; Pred. No. 46;
Matches 11; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 6 uncuuuuuuagccccaagggcu 29
Db 37671 TTTTCTTTTGAAGTCCCAAGGCT 37648

RESULT 6
US-08-781-891-209
Sequence 209, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 60.0%; Score 17.4; DB 3; Length 51259;
Best Local Similarity 48.1%; Pred. No. 54;
Matches 13; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 3 agauuuuuuuuagccccaagggcu 29
Db 26015 AGTTCTTTTGGAGACCTTAGACT 26041

RESULT 7
US-08-781-891-209/c
Sequence 209, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 60.0%; Score 17.4; DB 3; Length 51259;
Best Local Similarity 48.1%; Pred. No. 54;
Matches 13; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 3 agauuuuuuuuagccccaagggcu 29

Db 10496 AGTTCCTTTTGTGACACTTAGACT 10470

RESULT 8

US-08-222-619-1/c
Sequence 1, Application US/08222619
Patent No. 5652352
GENERAL INFORMATION:
APPLICANT: Lichenstein, Henri
APPLICANT: Lyons, David
APPLICANT: Wurfel, Mark
APPLICANT: Wright, Samuel
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,619
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2287 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 318..2117
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 381..2114
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 318..380
-08-222-619-1

Query Match 59.3%; Score 17.2; DB 1; Length 2287;
Best Local Similarity 54.5%; Pred. No. 40;
Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 agauucuuuuuuuagcccccaa 24

Db 175 AGACTCTTTTGTGACACCAA 154

RESULT 9

US-08-221-767-23/c
Sequence 23, Application US/08221767
Patent No. 6268212
GENERAL INFORMATION:
APPLICANT: Simonet, William S.
APPLICANT: Lichenstein, Henri S.
APPLICANT: Lyons, David E.
TITLE OF INVENTION: Tissue Specific Transgene Expression
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc., U.S. Patent Operations/NAO
STREET: 1840 DeHavilland Drive

CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,767
FILING DATE:
CLASSIFICATION: 800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2287 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 318..2117
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 381..2114
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 318..380
US-08-221-767-23

Query Match 59.3%; Score 17.2; DB 4; Length 2287;
Best Local Similarity 54.5%; Pred. No. 40;
Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 agauucuuuuuuuagcccccaa 24

Db 175 AGACTCTTTTGTGACACCAA 154

RESULT 10

PCT-US95-04075-1/c
Sequence 1, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2287 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 318..2117
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 381..2114
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 318..380
PCF-US95-04075-1

Query Match 59.3%; Score 17.2; DB 5; Length 2287;
Best Local Similarity 54.5%; Pred. No. 40;
Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 aagaucuuuuuuaagcccaaa 24
DB 175 AGACTCTTTTGTGACGCCAA 154

RESULT 11

US-08-928-926A-2/c
Sequence 2, Application US/08928926A

Patent No. 5925543

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Tang, Y. Tom

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: NADH DEHYDROGENASE B17 SUBUNIT

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,926A

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0384 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: TBLYN0101

CLONE: 44898

US-08-928-926A-2

Query Match 57.9%; Score 16.8; DB 2; Length 610;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 aagaucuuuuuuaagcccaagggc 28
DB 318 AGACTCTTTTGTATACCCCATGAC 291

RESULT 12

US-09-212-149-2/c

Sequence 2, Application US/09212149

Patent No. 6100036

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Tang, Y. Tom

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: NADH DEHYDROGENASE B17 SUBUNIT

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/212,149

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/928,926

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0384 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: TBLYN0101

CLONE: 44898

US-09-212-149-2

Query Match 57.9%; Score 16.8; DB 3; Length 610;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 aagaucuuuuuuaagcccaagggc 28
DB 318 AGACTCTTTTGTATACCCCATGAC 291

RESULT 13

US-08-998-416-900

Sequence 900, Application US/08998416

Patent No. 6239264

GENERAL INFORMATION:

APPLICANT: Philippson, Peter

APPLICANT: Pohlmann, Rainer

APPLICANT: Steiner, Sabine

APPLICANT: Mohr, Christine

APPLICANT: Wendland, Jurgen

APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPITII
NUMBER OF SEQUENCES: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6239264rtis Corporation
STREET: 3054 Cornwalls Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 900:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1561UP
US-08-998-416-900

Query Match 57.2%; Score 16.6; DB 4; Length 711;
Best Local Similarity 47.8%; Pred. No. 61;
Matches 11; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 6 uucuuuuuagcccaaggc 28
:::|||||:| |||||
168 TTCTTTTCTTTTCACCAAGGC 190

RESULT 14
US-08-834-306-17/c
Sequence 17, Application US/08834306
Patent No. 6054135
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,306
FILING DATE: 15-APR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-834-306-17

Query Match 57.2%; Score 16.6; DB 3; Length 2446;
Best Local Similarity 56.5%; Pred. No. 75;
Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 6 uucuuuuuagcccaaggc 28
:::|||||:| |||||
DB 370 TTCTTCTGTACTCCCAAGGC 348

RESULT 15
US-08-993-674A-17/c
Sequence 17, Application US/08993674A
Patent No. 6228372
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Smith, John M.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,674A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-993-674A-17

Query Match	Score	DB 4	Length	2446;
57.28;				
56.58;				
56.00;				
55.50;				
55.00;				
54.50;				
54.00;				
53.50;				
53.00;				
52.50;				
52.00;				
51.50;				
51.00;				
50.50;				
50.00;				
49.50;				
49.00;				
48.50;				
48.00;				
47.50;				
47.00;				
46.50;				
46.00;				
45.50;				
45.00;				
44.50;				
44.00;				
43.50;				
43.00;				
42.50;				
42.00;				
41.50;				
41.00;				
40.50;				
40.00;				
39.50;				
39.00;				
38.50;				
38.00;				
37.50;				
37.00;				
36.50;				
36.00;				
35.50;				
35.00;				
34.50;				
34.00;				
33.50;				
33.00;				
32.50;				
32.00;				
31.50;				
31.00;				
30.50;				
30.00;				
29.50;				
29.00;				
28.50;				
28.00;				
27.50;				
27.00;				
26.50;				
26.00;				
25.50;				
25.00;				
24.50;				
24.00;				
23.50;				
23.00;				
22.50;				
22.00;				
21.50;				
21.00;				
20.50;				
20.00;				
19.50;				
19.00;				
18.50;				
18.00;				
17.50;				
17.00;				
16.50;				
16.00;				
15.50;				
15.00;				
14.50;				
14.00;				
13.50;				
13.00;				
12.50;				
12.00;				
11.50;				
11.00;				
10.50;				
10.00;				
9.50;				
9.00;				
8.50;				
8.00;				
7.50;				
7.00;				
6.50;				
6.00;				
5.50;				
5.00;				
4.50;				
4.00;				
3.50;				
3.00;				
2.50;				
2.00;				
1.50;				
1.00;				
0.50;				
0.00;				

Best Local Similarity	56.5%;	Pred. No. 75;			
Matches	13; Conservative	6; Mismatches	4; Indels	0; Gaps	0;

Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

6 ucuuuuuagccccaaggc 28

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~

Search completed: October 2, 2001, 05:01:52
Job time: 5319 sec

Job time: 5319 sec

THIS PAGE BLANK (USPTO)